PRIME EDITOR VARIANTS, CONSTRUCTS, AND METHODS FOR ENHANCING PRIME EDITING EFFICIENCY AND PRECISION

GOVERNMENT SUPPORT

[1] This invention was made with government support under Grant Nos. AI142756, AI150551, HG009490, EB022376, EB031172, GM118062, CA072720, GM138167, U01AI142756, RM1HG009490, R01EB022376, and R35GM118062 awarded by the National Institutes of Health, and Grant No. HR0011-17-2-0049 awarded by the Department of Defense. The government has certain rights in the invention.

RELATED APPLICATIONS

[2] This application claims the priority under 35 U.S.C. § 119(e) to U.S. Provisional Application U.S.S.N. 63/255,897, filed October 14, 2021, U.S. Provisional Application U.S.S.N. 63/231,230, filed August 9, 2021, U.S. Provisional Application U.S.S.N. 63/194,913, filed May 28, 2021, U.S. Provisional Application U.S.S.N. 63/194,865, filed May 28, 2021, U.S. Provisional Application U.S.S.N. 63/176,202, filed April 16, 2021, and U.S. Provisional Application U.S.S.N. 63/136,194, filed January 11, 2021, each of which is incorporated herein by reference.

INCORPORATION BY REFERENCE

[3] In addition, this application refers to and incorporates by reference the entire contents of each of the following patent applications directed to prime editing previously filed by one or more of the present inventors: U.S. Provisional Application U.S.S.N. 62/820,813, filed March 19, 2019; U.S. Provisional Application U.S.S.N. 62/858,958, filed June 7, 2019; U.S. Provisional Application U.S.S.N. 62/889,996, filed August 21, 2019; U.S. Provisional Application U.S.S.N. 62/922,654, filed August 21, 2019; U.S. Provisional Application U.S.S.N. 62/913,553, filed October 10, 2019; U.S. Provisional Application U.S.S.N. 62/973,558, filed October 10, 2019; U.S. Provisional Application U.S.S.N. 62/931,195, filed November 5, 2019; U.S. Provisional Application U.S.S.N. 62/944,231, filed December 5, 2019; U.S. Provisional Application U.S.S.N. 62/974,537, filed December 5, 2019; U.S. Provisional Application U.S.S.N.

62/991,069, filed March 17, 2020; U.S. Provisional Application U.S.S.N. 63/100,548, filed March 17, 2020; International PCT Application No. PCT/US2020/023721, filed March 19, 2020; International PCT Application No. PCT/US2020/023553, filed March 19, 2020; International PCT Application No. PCT/US2020/023583, filed March 19, 2020; International PCT Application No. PCT/US2020/023730, filed March 19, 2020; International PCT Application No. PCT/US2020/023713, filed March 19, 2020; International PCT Application No. PCT/US2020/023712, filed March 19, 2020; International PCT Application No. PCT/US2020/023727, filed March 19, 2020; International PCT Application No. PCT/US2020/023724, filed March 19, 2020; International PCT Application No. PCT/US2020/023725, filed March 19, 2020; International PCT Application No. PCT/US2020/023728, filed March 19, 2020; International PCT Application No. PCT/US2020/023732, filed March 19, 2020; and International PCT Application No. PCT/US2020/023723, filed March 19, 2020.

BACKGROUND OF THE INVENTION

[4] The recent development of prime editing enables the insertion, deletion, and/or replacement of genomic DNA sequences without requiring error-prone double-strand DNA breaks. See Anzalone *et al.*, "Search-and-replace genome editing without double-strand breaks or donor DNA," *Nature*, 2019, Vol.576, pp. 149-157, the contents of which are incorporated herein by reference. Prime editing uses an engineered Cas9 nickase–reverse transcriptase fusion protein (e.g., PE1 or PE2) paired with an engineered prime editing guide RNA (pegRNA) that not only directs Cas9 to a target genomic site, but also which encodes the information for installing the desired edit. Without wishing to be bound by any particular theory, prime editing proceeds through a presumed multi-step editing process: 1) the Cas9 domain binds and nicks the target genomic DNA site, which is specified by the pegRNA's spacer sequence; 2) the reverse transcriptase domain uses the nicked genomic DNA as a primer to initiate the synthesis of an edited DNA strand using an engineered extension on the pegRNA as a template for reverse transcription–this generates a single-stranded 3′ flap containing the edited DNA sequence; 3) cellular DNA repair resolves the 3′ flap intermediate by the displacement of a 5′ flap species that occurs via invasion by the edited 3′ flap, excision of the 5′ flap containing the original DNA

sequence, and ligation of the new 3′ flap to incorporate the edited DNA strand, forming a heteroduplex of one edited and one unedited strand; and 4) cellular DNA repair replaces the unedited strand within the heteroduplex using the edited strand as a template for repair, completing the editing process.

[5] Since 2019, prime editing has been applied to introduce genetic changes in a wide variety of cells and/or organisms. Given its rapid adoption, prime editing represents a powerful tool for genomic editing. Despite its versatility and wide-scale use, prime editing efficiency can vary widely across different edit classes, target loci, and cell types (Anzalone *et al.*, 2019). Thus, modifications to prime editing systems which result in increasing the specificity and/or efficiency of the prime editing process would significantly help advance the art. In particular, modifications that facilitate more efficient incorporation of the edited DNA strand synthesized by the prime editor into the target genomic site are desirable. It is also desirable to reduce the frequency of indel byproducts that can form as a result of prime editing. Such further modifications to prime editing would advance the art.

SUMMARY OF THE INVENTION

[6] In one aspect, the present disclosure relates to the observation that the efficiency and/or specificity of prime editing is impacted by a cell's own DNA mismatch repair (MMR) DNA repair pathway. MMR is a multi-factor pathway that is involved in correcting basepair mismatches and insertion/deletion mispairs generated during DNA replication and recombination. As described herein, the inventors developed a novel genetic screening method—referred to in one embodiment as "pooled CRISPRi screen for prime editing outcomes"—which led to the identification of various genetic determinates, including MMR, as affecting the efficiency and/or specificity of prime editing. Accordingly, in one aspect, the present disclosure provides novel prime editing systems comprising a means for inhibiting and/or evade the effects of MMR, thereby increasing the efficiency and/or specificity of prime editing. In one embodiment, the disclosure provides a prime editing system that comprises an MMR-inhibiting protein, such as, but not limited to, a dominant negative variant of an MMR protein, such as a dominant negative MLH1 protein (i.e., "MLH1dn"). In another embodiment, the prime editing system comprises the installation of one or more silent mutations nearby an intended edit,

thereby allowing the intended edit from evading MMR recognition, even in the absence of an MMR-inhibiting protein, such as an MLH1dn. In another aspect, the disclosure provides a novel genetic screen for identifying genetic determinants, such as MMR, that impact the efficiency and/or specificity of prime editing. In still further aspects, the disclosure provides nucleic acid constructs encoding the improved prime editing systems described herein. The disclosure in other aspects also provides vectors (e.g., AAV or lentivirus vectors) comprising nucleic acids encoding the improved prime editing system described herein. In still other aspects, the disclosure provides cells comprising the improved prime editing systems described herein. The disclosure also provides in other aspects the components of the genetic screens, including nucleic acid and/or vector constructs, guide RNA, pegRNAs, cells (e.g., CRISPRi cells), and other reagents and/or materials for conducting the herein disclosed genetic screens. In still other aspects, the disclosure provides compositions and kits, e.g., pharmaceutical compositions, comprising the improved prime editing system described herein and which are capable of being administered to a cell, tissue, or organism by any suitable means, such as by gene therapy, mRNA delivery, virus-like particle delivery, or ribonucleoprotein (RNP) delivery. In yet another aspect, the present disclosure provides methods of using the improved prime editing system to install one or more edits in a target nucleic acid molecule, e.g., a genomic locus. In still another aspect, the present disclosure provides methods of treating a disease or disorder using the improved prime editing system to correct or otherwise repair one or more genetic changes (e.g., a single nucleotide polymorphism) in a target nucleic acid molecule, e.g., a genomic locus comprising one or more disease-causing mutations.

[7] Thus, in various aspects, the present disclosure describes an improved and modified approach to prime editing that comprises inhibiting the DNA mismatch repair (MMR) system during prime editing. The inventors have surprisingly found that the editing efficiency of prime editing may be significantly increased (e.g., at least a 2-fold increase, at least a 3-fold increase, at least a 4-fold increase, at least a 5-fold increase, at least a 6-fold increase, at least a 7-fold increase, at least an 8-fold increase, at least a 9-fold increase, at least a 10-fold increase, or more) when one or more functions of the DNA mismatch repair (MMR) system are inhibited, blocked, or otherwise inactivated during prime editing (such as using the MLH1dn inhibitor of MMR). In addition, the inventors have surprisingly found that the frequency of indel formation resulting from prime editing may be significantly decreased (e.g., about a 2-fold decrease, about a 3-fold

decrease, about a 4-fold decrease, about a 5-fold decrease, about a 6-fold decrease, about a 7-fold decrease, about a 8-fold decrease, about a 9-fold decrease, or about a 10-fold decrease or lower) when one or more functions of the DNA mismatch repair (MMR) system are inhibited, blocked, or otherwise inactivated during prime editing.

[8] The present disclosure also describes in other embodiments an improved and modified approach to prime editing that comprises evading the DNA mismatch repair (MMR) system during prime editing. The inventors have surprisingly found that the editing efficiency of prime editing may be significantly increased (e.g., at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold increased) when one or more silent mutations are installed nearby a desired site for installing a genetic change by prime editing, in the presence or absence of an inhibitor of MMR. In addition, the inventors have surprisingly found that the frequency of indel formation resulting from prime editing may be significantly decreased (e.g., at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-

fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold decreased) when one or more silent mutations are installed nearby a desired site for installing a genetic change by prime editing, in the presence or absence of an inhibitor of MMR.

[9] In some embodiments, the disclosure describes an improved prime editing system referred to herein as "PE4," which includes PE2 plus an MLH1 dominant negative protein (*e.g.*, wild-type MLH1 with amino acids 754-756 truncated as described further herein). In certain embodiments, the MLH1dn is expressed *in trans* in a cell comprising the PE2 fusion protein. The MLH1dn and the PE2 may be provided together or separate, e.g., by delivery on separate plasmids, separate vectors (e.g., AAV or lentivirus vectors), separate vector-like particles, separate ribonucleoprotein complexes (RNPs), or by delivery on the same plasmids, same vectors (e.g., AAV or lentivirus vectors), same vector-like particles, same ribonucleoprotein complexes (RNPs). In other embodiments, the MLH1dn may be fused to PE2 or otherwise associated with, coupled, or joined to PE2 such that they are co-delivered.

[10] In other embodiments, the disclosure describes an improved prime editing system referred to as "PE5," which includes PE3 (which is PE2 plus a second-strand nicking guide RNA) plus an MLH1 dominant negative protein (*e.g.*, wild-type MLH1 with amino acids 754-756 truncated as described further herein). In certain embodiments, the MLH1dn is expressed *in trans* in a cell comprising the PE3 prime editor. The MLH1dn and the PE3 may be provide together or separate, e.g., by delivery on separate plasmids, separate vectors (e.g., AAV or lentivirus vectors), separate vector-like particles, separate ribonucleoprotein complexes (RNPs), or by delivery on the same plasmid, same vector (e.g., AAV or lentivirus vectors), same vector-like particles, same ribonucleoprotein complexes (RNPs). In other embodiments, the MLH1dn

may be fused to PE3 or otherwise associated with, coupled, or joined to PE3 such that they are co-delivered.

[11] In other aspects, the present disclosure describes an optimized PE2 prime editor architecture referred to herein as "PEmax." PEmax is a modified form of PE2 which comprises modified reverse transcriptase codon usage, SpCas9 mutations, NLS sequences, and is described in FIG. 54B. Specifically, PEmax refers to a PE complex comprising a fusion protein comprising Cas9 (R221K N394K H840A) and a variant MMLV RT pentamutant (D200N T306K W313F T330P L603W) having the following structure: [bipartite NLS]-[Cas9(R221K)(N394K)(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)]-[bipartite NLS]-[NLS] + a desired PEgRNA, wherein the PE fusion has the amino acid sequence of SEQ ID NO: 99, which is shown as follows:

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNT**
DRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADL
RLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ
LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIK
RYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI
LEKMDGTEELLVKLKREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKD
NREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI
ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKA
IVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDK
DFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGW
GRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQ
GDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE
LDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYW
RQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRM
NTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT
ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITL
ANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKE
SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKEL
LGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGEL
QKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFS
KRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSSGGSKRTADGSEFESPKKK
RKVSGGSSGGSTLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAP
LIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTND
YRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPL
FAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLL

LAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEAR
KETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKA
YQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDP
VAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMT
HYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDA
DHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEG
KKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCP
GHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP*SGGS*KRTADGSEFESPKK
KRKV*GSGPAAKRVKLD* (SEQ ID NO: 99)

KEY:

BIPARTITE SV40 NUCLEAR LOCALIZATION SEQUENCE (NLS) TOP: (SEQ ID NO: 101), BOTTOM: (SEQ ID NO: 140)

CAS9(R221K N39K H840A) (SEQ ID NO: 104)

SGGSX2-BIPARTITE SV40NLS-SGGSX2 LINKER (SEQ ID NO: 105)

M-MLV reverse transcriptase (D200N T306K W313F T330P L603W) (SEQ ID NO: 98)

Other linker sequence (SEQ ID NO: 122)

Other linker sequence (SEQ ID NO: 106)

c-Myc NLS PAAKRVKLD (SEQ ID NO: 135)

[12] In some embodiments, the PE4 may be modified to substitute the PE2 fusion protein with PEmax. In such cases, the modified prime editing system may be referred to as "PE4max."

[13] In some embodiments, the PE5 may be modified to substitute the PE3 prime editor with PEmax. In such cases, the modified prime editing system may be referred to as "PE5max" and includes a second stranding nicking guide RNA.

[14] The inventors developed prime editing which enables the insertion, deletion, and/or replacement of genomic DNA sequences without requiring error-prone double-strand DNA breaks. The present disclosure now provides an improved method of prime editing involving the blocking, inhibiting, evading, or inactivation of the MMR pathway (e.g., by inhibiting, blocking, or inactivating an MMR pathway protein, including MLH1) during prime editing, whereby doing so surprisingly results in increased editing efficiency and reduced indel formation. As used herein, "during" prime editing can embrace any suitable sequence of events, such that the prime editing step can be applied before, at the same time, or after the step of blocking, inhibiting, evading, or inactivating the MMR pathway (e.g., by targeting the inhibition of MLH1).

[15] In various aspects and without wishing to be bound by any particular theory, prime editing uses an engineered Cas9 nickase–reverse transcriptase fusion protein (e.g., PE1 or PE2) paired with an engineered prime editing guide RNA (pegRNA) that both directs Cas9 to the target genomic site and encodes the information for installing the desired edit. Prime editing proceeds through a multi-step editing process: 1) the Cas9 domain binds and nicks the target genomic DNA site, which is specified by the pegRNA's spacer sequence; 2) the reverse transcriptase domain uses the nicked genomic DNA as a primer to initiate the synthesis of an edited DNA strand using an engineered extension on the pegRNA as a template for reverse transcription–this generates a single-stranded 3′ flap containing the edited DNA sequence; 3) cellular DNA repair resolves the 3′ flap intermediate by the displacement of a 5' flap species that occurs via invasion by the edited 3′ flap, excision of the 5′ flap containing the original DNA sequence, and ligation of the new 3′ flap to incorporate the edited DNA strand, forming a heteroduplex of one edited and one unedited strand; and 4) cellular DNA repair replaces the unedited strand within the heteroduplex using the edited strand as a template for repair, completing the editing process.

[16] Efficient incorporation of the desired edit requires that the newly synthesized 3′ flap contains a portion of sequence that is homologous to the genomic DNA site. This homology enables the edited 3′ flap to compete with the endogenous DNA strand (the corresponding 5' flap) for incorporation into the DNA duplex. Because the edited 3' flap will contain less sequence homology than the endogenous 5′ flap, the competition is expected to favor the 5′ flap strand. Thus, a potential limiting factor in the efficiency of prime editing may be the failure of the 3′ flap, which contains the edit, to effectively invade and displace the 5′ flap strand. Moreover, successful 3′ flap invasion and removal of the 5′ flap only incorporates the edit on one strand of the double-stranded DNA genome. Permanent installation of the edit requires cellular DNA repair to replace the unedited complementary DNA strand using the edited strand as a template. While the cell can be made to favor replacement of the unedited strand over the edited strand (step 4 above) by the introduction of a nick in the unedited strand adjacent to the edit using a secondary sgRNA (i.e., the PE3 system), this process still relies on a second stage of DNA repair.

[17] This disclosure describes a modified approach to prime editing that comprises additionally inhibiting, blocking, or otherwise inactivating the DNA mismatch repair (MMR)

system. In certain embodiments, the DNA mismatch repair (MMR) system can be inhibited, blocked, or otherwise inactivating one or more proteins of the MMR system, including, but not limited to MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. The disclosure contemplates any suitable means by which to inhibit, block, or otherwise inactivate the DNA mismatch repair (MMR) system, including, but not limited to inactivating one or more critical proteins of the MMR system at the genetic level, e.g., by introducing one or more mutations in the genes encoding a protein of the MMR system, e.g., MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

[18] Thus, in one aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating the DNA mismatch repair (MMR) system.

[19] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating a protein of the MMR system, e.g., MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

[20] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating MLH1 or variant thereof.

[21] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating PMS2 (or MutL alpha) or variant thereof.

[22] In yet another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating PMS1 (or MutL beta) or variant thereof.

[23] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating MLH3 (or MutL gamma) or variant thereof.

[24] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating MutS alpha (MSH2-MSH6) or variant thereof.

[25] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating MSH2 or variant thereof.

[26] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating MSH6 or variant thereof.

[27] In yet another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating PCNA or variant thereof.

[28] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating RFC or variant thereof.

[29] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating EXO1 or variant thereof.

[30] In yet another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating POLδ or variant thereof.

[31] Thus, in one aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, evading, or otherwise inactivating the DNA mismatch repair (MMR) system.

[32] In another aspect, the disclosure provides a method for evading MMR by installing one or more silent mutations nearby an intended edit, resulting in the evading of MMR and thereby improving editing efficiency of prime editing. In various embodiments, the number of silent mutations installed can be one, or two, or three, or four, or five, or six, or seven, or eight, or nine, or ten, or eleven, or twelve, or thirteen, or fourteen, or fifteen, or sixteen, or seventeen, or eighteen, or nineteen, or twenty or more. The one more silent mutations may be located upstream or downstream (or a combination if multiple silent mutations are involved) of the

intended edit site, on the same or opposite strand of DNA as the intended edit site (or a combination if multiple silent mutations are involved). The silent mutations may be located upstream or downstream of the intended edit by 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, or more nucleotide positions away from the intended edit. In various embodiments, the method of evading by silent mutation installation results in a significant increase in editing efficiency of prime editing (e.g., at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold increased) when one or more silent mutations are installed nearby a desired site for installing a genetic change by prime editing, in the presence or absence of an inhibitor of MMR. In various embodiments, the method of evading MMR by silent mutation installation results in a significant decrease in the frequency of indel formation of prime editing (e.g., at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-

fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold decrease) when one or more silent mutations are installed nearby a desired site for installing a genetic change by prime editing, in the presence or absence of an inhibitor of MMR.

[33] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of the MMR system, e.g., an inhibitor of one or more of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, or PCNA. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an antibody, e.g., a neutralizing antibody. In still other embodiments, the inhibitor can be a variant of an MMR protein (*e.g.*, a variant encoded by a dominant negative mutant of the gene encoding the MMR protein that adversely affects the function or expression of the normal wild type MMR protein, also referred to herein as a "dominant negative mutant," "dominant negative variant," or a "dominant negative protein," e.g., a "dominant negative MMR protein"). In some embodiments, the inhibitor is a dominant negative variant of an MMR protein that inhibits the activity of a wild type MMR protein. For example, the inhibitor can be an MLH1 protein variant (*e.g.*, a dominant negative mutant) of one or more of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, or PCNA, e.g., a dominant negative mutant of MLH1. In still other embodiments, the inhibitor can be targeted at the level of transcription, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, or PCNA. In yet other embodiments, the step of "contacting a target nucleotide

molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell an mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[34] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MLH1 or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MLH1. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-MLH1 antibody, e.g., a neutralizing antibody that inactivates MLH1. In still other embodiments, the inhibitor can be a dominant negative mutant of MLH1. In still other embodiments, the inhibitor can be targeted at the level of transcription of MLH1, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MLH1. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell an mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[35] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PMS2 (or MutL alpha) or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of PMS2 (or MutL alpha). In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-PMS2 (or MutL alpha) antibody, e.g., a neutralizing antibody that inactivates PMS2 (or MutL alpha). In still other embodiments, the inhibitor can be a dominant

negative mutant of PMS2 (or MutL alpha). In still other embodiments, the inhibitor can be targeted at the level of transcription of PMS2 (or MutL alpha), e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding ML PMS2 (or MutL alpha). In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[36] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PMS1 (or MutL beta) or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of PMS1 (or MutL beta). In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-PMS1 (or MutL beta) antibody, e.g., a neutralizing antibody that inactivates PMS1 (or MutL beta). In still other embodiments, the inhibitor can be a dominant negative mutant of PMS1 (or MutL beta). In still other embodiments, the inhibitor can be targeted at the level of transcription of PMS1 (or MutL beta), e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding PMS1 (or MutL beta). In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[37] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MLH3 (or MutL gamma) or variant thereof. In another aspect, the present disclosure

provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MLH3 (or MutL gamma). In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-MLH3 (or MutL gamma) antibody, e.g., a neutralizing antibody that inactivates MLH3 (or MutL gamma). In still other embodiments, the inhibitor can be a dominant negative mutant of MLH3 (or MutL gamma). In still other embodiments, the inhibitor can be targeted at the level of transcription of P MLH3 (or MutL gamma), e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MLH3 (or MutL gamma). In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[38] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MutS alpha (MSH2-MSH6) or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MutS alpha (MSH2-MSH6). In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-MutS alpha (MSH2-MSH6) antibody, e.g., a neutralizing antibody that inactivates MutS alpha (MSH2-MSH6). In still other embodiments, the inhibitor can be a dominant negative mutant of MutS alpha (MSH2-MSH6). In still other embodiments, the inhibitor can be targeted at the level of transcription of MutS alpha (MSH2-MSH6), e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MutS alpha (MSH2-MSH6). In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or

lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[39] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MSH2 or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MSH2. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti- MSH2 antibody, e.g., a neutralizing antibody that inactivates MSH2. In still other embodiments, the inhibitor can be a dominant negative mutant of MSH2. In still other embodiments, the inhibitor can be targeted at the level of transcription of MSH2, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MSH2. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[40] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MSH6 or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MSH6. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-MSH6 antibody, e.g., a neutralizing antibody that inactivates MSH6. In still other embodiments, the inhibitor can be a dominant negative mutant of MSH6. In still other embodiments, the inhibitor can be targeted at the level of transcription of MSH6, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MSH6. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA

that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[41] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PCNA or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of PCNA. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti- PCNA antibody, e.g., a neutralizing antibody that inactivates PCNA. In still other embodiments, the inhibitor can be a dominant negative mutant of PCNA. In still other embodiments, the inhibitor can be targeted at the level of transcription of PCNA, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding PCNA. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[42] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating RFC or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of RFC. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-RFC antibody, e.g., a neutralizing antibody that inactivates RFC. In still other embodiments, the inhibitor can be a dominant negative mutant of RFC. In still other embodiments, the inhibitor can be targeted at the level of transcription of RFC, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding RFC. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery

system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[43] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating EXO1 or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of EXO1. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-EXO1 antibody, e.g., a neutralizing antibody that inactivates EXO1. In still other embodiments, the inhibitor can be a dominant negative mutant of EXO1. In still other embodiments, the inhibitor can be targeted at the level of transcription of EXO1, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding EXO1. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[44] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating POLδ or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of POLδ. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-POLδ antibody, e.g., a neutralizing antibody that inactivates POLδ. In still other embodiments, the inhibitor can be a dominant negative mutant of POLδ. In still other embodiments, the inhibitor can be targeted at the level of transcription of POLδ, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding POLδ. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly

to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[45] In one aspect, the present disclosure provides methods for editing a nucleic acid molecule by prime editing. In some embodiments, the method comprises contacting a nucleic acid molecule with a prime editor, a pegRNA, and an inhibitor of the DNA mismatch repair pathway, thereby installing one or more modifications to the nucleic acid molecule at a target site.

[46] The method may increase the efficiency of prime editing and/or decrease the frequency of indel formation. In some embodiments, the prime editing efficiency is increased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold in the presence of the inhibitor of the DNA mismatch repair pathway. In some embodiments, the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold in the presence of the inhibitor of the DNA mismatch repair pathway.

[47] In some embodiments, the inhibitor of the DNA mismatch repair pathway inhibits one or more proteins of the DNA mismatch repair pathway. In some embodiments, the one or more proteins is selected from the group consisting of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. In certain embodiments, the one or more proteins is MLH1. In some embodiments, MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

[48] The inhibitor utilized in the method may be an antibody, a small molecule, a small interfering RNA (siRNA), a small non-coding microRNA, or a dominant negative variant of an

MMR protein that inhibits the activity of a wild type MMR protein (*e.g.*, a dominant negative variant of MLH1). In certain embodiments, the inhibitor is an antibody that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a small molecule that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In certain embodiments, the inhibitor is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

[49] In some embodiments, the dominant negative variant is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.

[50] The prime editors utilized in the methods of the present disclosure may comprise multiple components. In some embodiments, the prime editor comprises a napDNAbp and a polymerase. In some embodiments, the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof. In certain embodiments, the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, Cas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute and optionally has a nickase activity. In certain embodiments, the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, or 99 (PEmax) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2, 4-67, or 99 (PEmax). In certain embodiments, the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 37 (e.g., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2. In some embodiments, the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase. In some embodiments, the

polymerase is a reverse transcriptase. In certain embodiments, the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98.

[51] The napDNAbp and the polymerase of the prime editor may be joined together to form a fusion protein. In some embodiments, the napDNAbp and the polymerase of the prime editor are joined by a linker to form a fusion protein. In certain embodiments, the linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, or 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, or 118-131. In some embodiments, the linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

[52] The components used in the method (*e.g.*, the prime editor, the pegRNA, and/or the inhibitor of the DNA mismatch repair pathway) may be encoded on a DNA vector. In some embodiments, the prime editor, the pegRNA, and the inhibitor of the DNA mismatch repair pathway are encoded on one or more DNA vectors. In certain embodiments, the one or more DNA vectors comprise AAV or lentivirus DNA vectors. In some embodiments, the AAV vector is serotype 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

[53] The prime editors utilized in the presently disclosed methods may also be further joined to additional components. In some embodiments, the prime editor as a fusion protein is further joined by a second linker to the inhibitor of the DNA mismatch repair pathway. In certain embodiments, the second linker is a self-hydrolyzing linker. In certain embodiments, the second linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or 233-236, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131, or 233-236. In some embodiments, the second linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

[54] In some embodiments, the one or more modifications to the nucleic acid molecule installed at the target site comprise one or more transitions, one or more transversions, one or more insertions, one or more deletions, or one more inversions. In certain embodiments, the one

or more transitions are selected from the group consisting of: (a) T to C; (b) A to G; (c) C to T; and (d) G to A. In certain embodiments, the one or more transversions are selected from the group consisting of: (a) T to A; (b) T to G; (c) C to G; (d) C to A; (e) A to T; (f) A to C; (g) G to C; and (h) G to T. In certain embodiments, the one or more modifications comprises changing (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair to a C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair. In some embodiments, the one or more modifications comprises an insertion or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides.

[55] The methods of the present disclosure may be used for making corrections to one or more disease-associated genes. In some embodiments, the one or more modifications comprises a correction to a disease-associated gene. In certain embodiments, the disease-associated gene is associated with a polygenic disorder selected from the group consisting of: heart disease; high blood pressure; Alzheimer's disease; arthritis; diabetes; cancer; and obesity. In certain embodiments, the disease-associated gene is associated with a monogenic disorder selected from the group consisting of: Adenosine Deaminase (ADA) Deficiency; Alpha-1 Antitrypsin Deficiency; Cystic Fibrosis; Duchenne Muscular Dystrophy; Galactosemia; Hemochromatosis; Huntington's Disease; Maple Syrup Urine Disease; Marfan Syndrome; Neurofibromatosis Type 1; Pachyonychia Congenita; Phenylketonuria; Severe Combined Immunodeficiency; Sickle Cell Disease; Smith-Lemli-Opitz Syndrome; a trinucleotide repeat disorder; a prion disease; and Tay-Sachs Disease.

[56] In another aspect, the present disclosure provides compositions for editing a nucleic acid molecule by prime editing. In some embodiments, the composition comprises a prime editor, a pegRNA, and an inhibitor of the DNA mismatch repair pathway, wherein the composition is capable of installing one or more modifications to the nucleic acid molecule at a target site.

[57] The composition may increase the efficiency of prime editing and/or decrease the frequency of indel formation. In some embodiments, the prime editing efficiency is increased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least

7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold in the presence of the inhibitor of the DNA mismatch repair pathway. In some embodiments, the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold in the presence of the inhibitor of the DNA mismatch repair pathway.

[58] In some embodiments, the inhibitor of the DNA mismatch repair pathway inhibits one or more proteins of the DNA mismatch repair pathway. In some embodiments, the one or more proteins is selected from the group consisting of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. In certain embodiments, the one or more proteins is MLH1. In some embodiments, MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

[59] The inhibitor utilized in the composition may be an antibody, a small molecule, a small interfering RNA (siRNA), a small non-coding microRNA, or a dominant negative variant of an MMR protein that inhibits the activity of a wild type MMR protein (*e.g.*, a dominant negative variant of MLH1). In certain embodiments, the inhibitor is an antibody that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a small molecule that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In certain embodiments, the inhibitor is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

[60] In some embodiments, the dominant negative variant is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO:

218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.

[61] The prime editors utilized in the compositions of the present disclosure comprise multiple components. In some embodiments, the prime editor comprises a napDNAbp and a polymerase. In some embodiments, the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof. In certain embodiments, the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, Cas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute and optionally has a nickase activity. In certain embodiments, the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, or 99 (PEmax) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2, 4-67, or 99 (PEmax). In certain embodiments, the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 37 (i.e., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2 or SEQ ID NO: 37. In some embodiments, the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase. In some embodiments, the polymerase is a reverse transcriptase. In certain embodiments, the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98.

[62] The napDNAbp and the polymerase of the prime editor may be joined together to form a fusion protein. In some embodiments, the napDNAbp and the polymerase of the prime editor are joined by a linker to form a fusion protein. In certain embodiments, the linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131. In some embodiments, the linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

[63] The components used in the compositions disclosed herein (*e.g.*, the prime editor, the pegRNA, and/or the inhibitor of the DNA mismatch repair pathway) may be encoded on a DNA vector. In some embodiments, the prime editor, the pegRNA, and the inhibitor of the DNA mismatch repair pathway are encoded on one or more DNA vectors. In certain embodiments, the one or more DNA vectors comprise AAV or lentivirus DNA vectors. In some embodiments, the AAV vector is serotype 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

[64] The prime editors utilized in the presently disclosed compositions may also be further joined to additional components. In some embodiments, the prime editor as a fusion protein is further joined by a second linker to the inhibitor of the DNA mismatch repair pathway. In certain embodiments, the second linker is a self-hydrolyzing linker. In certain embodiments, the second linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or 233-236, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131, or 233-236. In some embodiments, the second linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

[65] In some embodiments, the one or more modifications to the nucleic acid molecule installed at the target site comprise one or more transitions, one or more transversions, one or more insertions, one or more deletions, or one more inversions. In certain embodiments, the one or more transitions are selected from the group consisting of: (a) T to C; (b) A to G; (c) C to T; and (d) G to A. In certain embodiments, the one or more transversions are selected from the group consisting of: (a) T to A; (b) T to G; (c) C to G; (d) C to A; (e) A to T; (f) A to C; (g) G to C; and (h) G to T. In certain embodiments, the one or more modifications comprises changing (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair to a C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair. In some embodiments, the one or more modifications comprises an insertion or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides.

[66] The compositions of the present disclosure may be used for making corrections to one or more disease-associated genes. In some embodiments, the one or more modifications comprises a correction to a disease-associated gene. In certain embodiments, the disease-associated gene is associated with a polygenic disorder selected from the group consisting of: heart disease; high blood pressure; Alzheimer's disease; arthritis; diabetes; cancer; and obesity. In certain embodiments, the disease-associated gene is associated with a monogenic disorder selected from the group consisting of: Adenosine Deaminase (ADA) Deficiency; Alpha-1 Antitrypsin Deficiency; Cystic Fibrosis; Duchenne Muscular Dystrophy; Galactosemia; Hemochromatosis; Huntington's Disease; Maple Syrup Urine Disease; Marfan Syndrome; Neurofibromatosis Type 1; Pachyonychia Congenita; Phenylketonuria; Severe Combined Immunodeficiency; Sickle Cell Disease; Smith-Lemli-Opitz Syndrome; a trinucleotide repeat disorder; a prion disease; and Tay-Sachs Disease.

[67] In another aspect, this disclosure provides polynucleotides for editing a DNA target site by prime editing. In some embodiments, the polynucleotide comprises a nucleic acid sequence encoding a napDNAbp, a polymerase, and an inhibitor of the DNA mismatch repair pathway, wherein the napDNAbp and polymerase is capable in the presence of a pegRNA of installing one or more modifications in the DNA target site.

[68] The polynucleotide may increase the efficiency of prime editing and/or decrease the frequency of indel formation. In some embodiments, the prime editing efficiency is increased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold in the presence of the inhibitor of the DNA mismatch repair pathway. In some embodiments, the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold in the presence of the inhibitor of the DNA mismatch repair pathway.

[69] In some embodiments, the inhibitor of the DNA mismatch repair pathway inhibits one or more proteins of the DNA mismatch repair pathway. In some embodiments, the one or more proteins is selected from the group consisting of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL

beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. In certain embodiments, the one or more proteins is MLH1. In some embodiments, MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

[70] The inhibitor utilized in the polynucleotide may be an antibody, a small molecule, a small interfering RNA (siRNA), a small non-coding microRNA, or a dominant negative variant of an MMR protein that inhibits the activity of a wild type MMR protein (*e.g.*, a dominant negative variant of (MLH1). In certain embodiments, the inhibitor is an antibody that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a small molecule that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In certain embodiments, the inhibitor is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

[71] In some embodiments, the dominant negative variant is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.

[72] The prime editors utilized in the polynucleotides of the present disclosure comprise multiple components (*e.g.*, a napDNAbp and a polymerase). In some embodiments, the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof. In certain embodiments, the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, Cas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j

(CasΦ), and Argonaute and optionally has a nickase activity. In certain embodiments, the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, or 99 (PEmax) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2, 4-67, or 99 (PEmax). In certain embodiments, the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 37 (i.e., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2 or SEQ ID NO: 37. In some embodiments, the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase. In some embodiments, the polymerase is a reverse transcriptase. In certain embodiments, the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98.

[73] The napDNAbp and the polymerase of the prime editor may be joined together to form a fusion protein. In some embodiments, the napDNAbp and the polymerase of the prime editor are joined by a linker to form a fusion protein. In certain embodiments, the linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131. In some embodiments, the linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

[74] The polynucleotides disclosed herein may comprise vectors. In some embodiments, the polynucleotide is a DNA vector. In certain embodiments, the DNA vector is an AAV or lentivirus DNA vector. In some embodiments, the AAV vector is serotype 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

[75] The prime editors encoded by the presently disclosed polynucleotides may also be further joined to additional components. In some embodiments, the prime editor as a fusion protein is further joined by a second linker to the inhibitor of the DNA mismatch repair pathway. In certain embodiments, the second linker comprises a self-hydrolyzing linker. In certain embodiments, the second linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or 233-236, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131, or 233-236. In some embodiments, the

second linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

[76] In some embodiments, the one or more modifications to the nucleic acid molecule installed at the target site comprise one or more transitions, one or more transversions, one or more insertions, one or more deletions, or one more inversions. In certain embodiments, the one or more transitions are selected from the group consisting of: (a) T to C; (b) A to G; (c) C to T; and (d) G to A. In certain embodiments, the one or more transversions are selected from the group consisting of: (a) T to A; (b) T to G; (c) C to G; (d) C to A; (e) A to T; (f) A to C; (g) G to C; and (h) G to T. In certain embodiments, the one or more modifications comprises changing (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair to a C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair. In some embodiments, the one or more modifications comprises an insertion or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides.

[77] The polynucleotides of the present disclosure may be used for making corrections to one or more disease-associated genes. In some embodiments, the one or more modifications comprises a correction to a disease-associated gene. In certain embodiments, the disease-associated gene is associated with a polygenic disorder selected from the group consisting of: heart disease; high blood pressure; Alzheimer's disease; arthritis; diabetes; cancer; and obesity. In certain embodiments, the disease-associated gene is associated with a monogenic disorder selected from the group consisting of: Adenosine Deaminase (ADA) Deficiency; Alpha-1 Antitrypsin Deficiency; Cystic Fibrosis; Duchenne Muscular Dystrophy; Galactosemia; Hemochromatosis; Huntington's Disease; Maple Syrup Urine Disease; Marfan Syndrome; Neurofibromatosis Type 1; Pachyonychia Congenita; Phenylketonuria; Severe Combined Immunodeficiency; Sickle Cell Disease; Smith-Lemli-Opitz Syndrome; a trinucleotide repeat disorder; a prion disease; and Tay-Sachs Disease.

[78] In another aspect, the present disclosure provides cells. In some embodiments, the cell comprises any of the polynucleotides described herein.

[79] In another aspect, the present disclosure provides pharmaceutical compositions. In some embodiments, the pharmaceutical composition comprises any of the compositions disclosed herein. In some embodiments, the pharmaceutical composition comprises any of the compositions disclosed herein and a pharmaceutically acceptable excipient. In some embodiments, the pharmaceutical composition comprises any of the polynucleotides disclosed herein. In some embodiments, the pharmaceutical composition comprises any of the polynucleotides disclosed herein and a pharmaceutically acceptable excipient.

[80] In another aspect, the present disclosure provides kits. In some embodiments, the kit comprises any of the compositions disclosed herein, a pharmaceutical excipient, and instructions for editing a DNA target site by prime editing. In some embodiments, the kit comprises any of the polynucleotides disclosed herein, a pharmaceutical excipient, and instructions for editing a DNA target site by prime editing.

[81] The present disclosure also provides methods and pegRNAs for prime editing whereby correction by the MMR pathway of the alterations introduced into a target nucleic acid molecule is evaded, without the need to provide an inhibitor of the MMR pathway. Surprisingly, pegRNAs designed with consecutive nucleotide mismatches compared to the endogenous sequence of a target site on a target nucleic acid, for example, pegRNAs that have three or more consecutive mismatching nucleotides, can evade correction by the MMR pathway, resulting in an increase in prime editing efficiency and a decrease in the frequency of indel formation compared to the introduction of a single nucleotide mismatch using prime editing. In addition, insertions or deletions of consecutive nucleotides at the target site of the target nucleic acid, for example, insertions or deletions greater than 10 nucleotides in length, introduced by prime editing also evade correction by the MMR pathway, resulting in an increase in prime editing efficiency and a decrease in the frequency of indel formation compared to the introduction of an insertion or deletion of less than 10 nucleotides in length using prime editing.

[82] Thus, in another aspect, the present disclosure provides methods for editing a nucleic acid molecule by prime editing comprising contacting a nucleic acid molecule with a prime editor (*e.g.*, PE2, PE3, or any of the other prime editors described herein) and a pegRNA with a DNA synthesis template on its extension arm comprising three or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule. In some embodiments, at least one of the consecutive nucleotide mismatches results in an alteration

in the amino acid sequence of a protein expressed from the nucleic acid molecule, while at least one of the remaining nucleotide mismatches is a silent mutation. The silent mutations may be in coding regions of the target nucleic acid molecule (*i.e.*, in a part of a gene that encodes a protein), or the silent mutations may be in non-coding regions of the target nucleic acid molecule. In some embodiments, when the silent mutations are in a coding region, the silent mutations introduce into the nucleic acid molecule one or more alternate codons encoding the same amino acid as the unedited nucleic acid molecule. In some embodiments, when the silent mutations are in a non-coding region, the silent mutations are present in a region of the nucleic acid molecule that does not influence splicing, gene regulation, RNA lifetime, or other biological properties of the target site on the nucleic acid molecule.

[83] Any number of consecutive nucleotide mismatches compared to the sequence of the target site can be designed in the DNA synthesis template of a pegRNA to achieve the benefits of evading correction by the MMR pathway, and thereby increase prime editing efficiency and/or reduce indel formation. In some embodiments, the DNA synthesis template comprises at least three consecutive nucleotide mismatches compared to the sequence of the target site. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 consecutive nucleotide mismatches relative to the endogenous sequence of a target site in the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, or ten or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule. In certain embodiments, the use of three or more consecutive nucleotide mismatches results in an increase in prime editing efficiency by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a method using a pegRNA comprising a DNA synthesis template comprising only one consecutive nucleotide mismatch relative to the endogenous sequence of a target site on the nucleic acid molecule. In certain embodiments, the use of three or more consecutive nucleotide mismatches results in a decrease in the frequency of indel formation by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least

3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a method using a pegRNA comprising a DNA synthesis template comprising only one consecutive nucleotide mismatch relative to the endogenous sequence of a target site on the nucleic acid molecule.

[84] In another aspect, the present disclosure provides methods for editing a nucleic acid molecule by prime editing comprising contacting a nucleic acid molecule with a prime editor (*e.g.*, PE2, PE3, or any of the other prime editors described herein) and a pegRNA with a DNA synthesis template on its extension arm comprising an insertion or deletion of 10 or more contiguous nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule. In some embodiments, the DNA synthesis template of a pegRNA can be designed to introduce insertions or deletions greater than 3 nucleotides to avoid or reduce the impact of mismatch correction by the cellular MMR pathway, thereby improving prime editing efficiency. In some embodiments, the DNA synthesis template of the pegRNA is designed to introduce one or more insertions and/or deletions of 3, 4, 5, 6, 7, 8, 9, 10, or more contiguous nucleotides to avoid or reduce the impact of mismatch correction by the cellular MMR pathway, thereby improving prime editing efficiency. In some embodiments, insertions or deletions of any length greater than 10 contiguous nucleotides can be used to achieve the benefits of evading correction by the MMR pathway. In some embodiments, the DNA synthesis template comprises an insertion of 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides relative to the endogenous sequence of a target site on a nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template comprises a deletion of 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 contiguous nucleotides relative to the endogenous sequence of a target site on a nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template comprises an insertion of 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 contiguous nucleotides relative to the endogenous sequence of a target site on a nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template comprises a deletion of 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 contiguous nucleotides relative to the endogenous sequence of a target site on a nucleic acid molecule edited by prime

editing. In some embodiments, the DNA synthesis template comprises an insertion or deletion of 11 or more contiguous nucleotides, 12 or more contiguous nucleotides, 13 or more contiguous nucleotides, 14 or more contiguous nucleotides, 15 or more contiguous nucleotides, 16 or more contiguous nucleotides, 17 or more contiguous nucleotides, 18 or more contiguous nucleotides, 19 or more contiguous nucleotides, 20 or more contiguous nucleotides, 21 or more contiguous nucleotides, 22 or more contiguous nucleotides, 23 or more contiguous nucleotides, 24 or more contiguous nucleotides, or 25 or more contiguous nucleotides relative to a target site on a nucleic acid molecule. In certain embodiments, the DNA synthesis template comprises an insertion or deletion of 15 or more contiguous nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule.

[85] In some embodiments, prime editing with a pegRNA designed to introduce an insertion and/or deletion of multiple contiguous nucleotides, for example, three or more contiguous nucleotides, relative to the endogenous sequence of a target site results in an increase in prime editing efficiency compared to prime editing with a corresponding control pegRNA (e.g., a control pegRNA that does not introduce an insertion or deletion of three or more contiguous nucleotides) by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold. In some embodiments, prime editing with a pegRNA designed to introduce an insertion or deletion of 3, 4, 5, 6, 7, 8, 9, 10, or more contiguous nucleotides relative to the endogenous sequence of a target site results in an increase in prime editing efficiency relative to prime editing with a corresponding control pegRNA (e.g., a control pegRNA that does not introduce insertion or deletion of the three or more contiguous nucleotides) by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold. In some embodiments, making an insertion or deletion of 10 or more contiguous nucleotides results in an increase in prime editing efficiency by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-

fold relative to a method using a pegRNA comprising a DNA synthesis template comprising an insertion or deletion of fewer than 10 nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule. In some embodiments, making an insertion or deletion of 10 or more nucleotides results in a decrease in the frequency of indel formation by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a method using a pegRNA comprising a DNA synthesis template comprising an insertion or deletion of fewer than 10 nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule.

[86] In another aspect, the present disclosure also provides pegRNAs useful for editing a nucleic acid molecule by prime editing while evading correction by the MMR pathway of the alterations introduced into the nucleic acid molecule, thereby increasing prime editing efficiency and/or reducing indel formation. In some embodiments, the extension arm of the pegRNAs provided by the present disclosure comprise three or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule. In some embodiments, at least one of the three consecutive nucleotide mismatches relative to the endogenous sequence of the target site is a silent mutation. In some embodiments, at least one of the consecutive nucleotide mismatches results in an alteration in the amino acid sequence of a protein expressed from the target nucleic acid molecule, while at least one of the remaining nucleotide mismatches is a silent mutation. The silent mutations may be in coding regions of the target nucleic acid molecule (*i.e.*, in a part of a gene that encodes a protein), or the silent mutations may be in non-coding regions of the target nucleic acid molecule. In some embodiments, when the silent mutations are in a coding region, the silent mutations introduce into the nucleic acid molecule one or more alternate codons encoding the same amino acid as the unedited nucleic acid molecule. In some embodiments, when the silent mutations are in a non-coding region, the silent mutations are present in a region of the nucleic acid molecule that does not influence splicing, gene regulation, RNA lifetime, or other biological properties of the target site on the nucleic acid molecule.

[87] Any number of consecutive nucleotide mismatches of three or more can be incorporated into the extension arm of the pegRNAs described herein to achieve the benefits of evading

correction by the MMR pathway. In some embodiments, the DNA synthesis template of the extension arm of the pegRNA comprises 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm of the pegRNA comprises at least three consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm of the pegRNA comprises 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm of the pegRNA comprises 3, 4, 5, 6, 7, 8, 9, or 10 consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, or ten or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule. In certain embodiments, the presence of three or more consecutive nucleotide mismatches on the extension arm of the pegRNA results in an increase in prime editing efficiency by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a pegRNA comprising a DNA synthesis template comprising only one consecutive nucleotide mismatch relative to the endogenous sequence of a target site on the nucleic acid molecule. In certain embodiments, the use of three or more consecutive nucleotide mismatches results in a decrease in the frequency of indel formation by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a pegRNA comprising a DNA synthesis template comprising only one consecutive nucleotide mismatch relative to the endogenous sequence of a target site on the nucleic acid molecule.

[88] In another aspect, the present disclosure provides a prime editor system for site specific genome modification comprising (a) a prime editor comprising (i) a nucleic acid programmable DNA binding protein (napDNAbp) and (ii) a DNA polymerase, and (b) an inhibitor of the DNA mismatch repair pathway. In some embodiments, the inhibitor of the DNA mismatch repair pathway inhibits one or more proteins of the DNA mismatch repair pathway (*e.g.*, MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and/or PCNA). In some embodiments, the one or more proteins is MLH1. In certain embodiments, the MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

[89] Any inhibitor of the DNA mismatch repair pathway may be used in the systems described herein. In some embodiments, the inhibitor is an antibody that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a small molecule that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of one or more proteins of the DNA mismatch repair pathway. In some embodiments, the inhibitor is a dominant negative variant of an MMR protein that inhibits the activity of a wild type MMR protein (*e.g.*, a dominant negative variant of MLH1 that inhibits MLH1).

[90] In certain embodiments, the dominant negative variant used in the systems of the present disclosure is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLSSV40 (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLSSV40 MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223. The present disclosure also contemplates methods for performing prime editing on a nucleic acid molecule in a cell in which MMR activity is knocked

out entirely (*e.g.*, by knocking down one or more genes involved in the MMR pathway in the genome of the cell). Such methods provide the benefits of inhibiting MMR (*e.g.*, improved editing efficiency and decreased indel formation) without the need to provide an inhibitor of MMR. Thus, in another aspect, the present disclosure provides methods for editing a nucleic acid molecule by prime editing comprising: contacting a nucleic acid molecule with a prime editor and a pegRNA, thereby installing one or more modifications to the nucleic acid molecule at a target site, wherein the nucleic acid molecule is in a cell comprising a knockout of one or more genes involved in the DNA mismatch repair (MMR) pathway. In some embodiments, the method further comprises contacting the nucleic acid molecule with a second strand nicking gRNA. In certain embodiments, the prime editing efficiency is increased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a method performed in a cell that does not comprise a knockout of one or more genes involved in MMR. In certain embodiments, the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a method performed in a cell that does not comprise a knockout of one or more genes involved in MMR. In some embodiments, the one or more genes involved in MMR is selected from the group consisting of genes encoding the proteins MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. In certain embodiments, the one or more genes is the gene encoding MLH1 (*e.g.*, comprising an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204).

[91] In another aspect, the present disclosure provides methods for editing a nucleic acid molecule by prime editing comprising: contacting a nucleic acid molecule with a prime editor, a pegRNA, and an inhibitor of p53, thereby installing one or more modifications to the nucleic

acid molecule at a target site. In some embodiments, the method further comprises contacting the nucleic acid molecule with a second strand nicking gRNA.

[92] In some embodiments, the prime editing efficiency is increased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of p53. In some embodiments, the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at

least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of p53.

[93] In some embodiments, the inhibitor of p53 is a protein. In certain embodiments, the inhibitor of p53 is the protein i53. In some embodiments, the inhibitor of p53 is an antibody that inhibits the activity of p53. In some embodiments, the inhibitor of p53 is a small molecule that inhibits the activity of p53. In some embodiments, the inhibitor of p53 is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of p53.

[94] In another aspect, the present disclosure describes improved prime editor fusion proteins, including PEmax of SEQ ID NO: 99. The disclosure also contemplates fusion proteins having an amino acid sequence with a sequence identity of at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least up to 100% with SEQ ID NO: 99.

[95] The inventors have surprisingly found that the editing efficiency of prime editing may be significantly increased (e.g., 2-fold increase, 3-fold increase, 4-fold increase, 5-fold increase, 6-fold increase, 7-fold increase, 8-fold increase, 9-fold increase, or 10-fold increase or more) when one or more components of the canonical prime editor fusion protein (i.e., PE2) are modified. Modifications may include a modified amino acid sequence of one or more components (e.g., a Cas9 component, a reverse transcriptase component, or a linker).

[96] In other aspects, the present disclosure also provides compositions and pharmaceutical compositions comprising PEmax, methods of prime editing using PEmax, polynucleotides and vectors encoding PEmax, and kits and cells comprising PEmax.

[97] It should be appreciated that the foregoing concepts, and additional concepts discussed below, may be arranged in any suitable combination, as the present disclosure is not limited in this respect. Further, other advantages and novel features of the present disclosure will become apparent from the following detailed description of various non-limiting embodiments when considered in conjunction with the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

[98] The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure, which can be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

[99] **FIG. 1** provides a schematic showing that prime editing enables guide RNA-templated genomic manipulations. DNA prime editing intermediates capable of being repaired by cellular factors are shown in boxes.

[100] **FIG. 2** provides a schematic for a DNA repair CRISPRi screen for prime editing outcomes.

[101] **FIGs. 3A-3C** show optimization of prime editing efficiency at the target site. **FIG. 3A** provides a schematic for the optimization process. **FIG. 3B** shows percent reads with a specified modification at a target site in HeLa cells. **FIG. 3C** shows percent reads with a specified modification at a target site with blasticidin selection in HeLa cells.

[102] **FIGs. 4A-4B** show a prime editing CRISPRi screen with a DNA repair library. **FIG. 4A** provides a schematic of the screening process. **FIG. 4B** shows percent reads with a specified modification in bulk editing of post-screen HeLa cells.

[103] **FIGs. 5A-5B** show that the CRISPRi screen reveals that DNA mismatch repair limits prime editing efficiency. Knockdown of mismatch repair proteins (MSH2, MSH6, PMS2, and MLH1) improves the efficiency of PE2 by 3-fold and PE3 by 2-fold.

[104] **FIGS. 6A-6C** show that siRNA knockdown of MMR improves prime editing in HEK293T cells. Editing results at multiple endogenous loci validate the findings of the CRISPRi screen.

[105] **FIGS. 7A-7B** show that complete MMR knockout dramatically enhances prime editing. In the absence of MMR, PE2 editing efficiency is shown to match PE3 editing efficiency.

[106] **FIG. 8A** provides a schematic for the mechanism of mismatch repair (MMR). In the first step, MSH2:MSH6 (MutSα) binds the mismatch and recruits MLH1:PMS2 (MutLα). The DNA nick signals to MMR which strand to repair. In the second step, MutLα indiscriminately incises the nicked strand 5' and 3' of the mismatch. In the third step, EXO1 excises the mismatch from MutLα-generated nicks. In the fourth step, POLδ resynthesizes the excised strand, followed by LIG1 ligation.

[107] **FIG. 8B** provides yet another schematic for the mechanism of mismatch repair, MMR, in eukaryotic cells. The left side of the schematic depicts 5′ MMR. (A) The MutS homolog proteins (MSH, purple) MutSα (MSH2-MSH6), or MutSβ (MSH2-MSH3) recognize and bind a mismatch. RPA bound to single-strand DNA prevents EXO1 from accessing and degrading DNA. (B) In the sliding clamp model, MutSα/β at a mismatch binds ATP and undergoes nucleotide switch activation, becoming a sliding clamp that diffuses along the DNA. Multiple MSH clamps are loaded at a single mismatch. The interaction of EXO1 with MSH sliding clamps overcomes the RPA barrier and activates EXO1 for 5′ to 3′ excision from the 5′ nick. MutL homolog proteins (MLH) (MutLα is ScMlh1-Pms1 or HsMLH1-PMS2) bind ATP and may interact with MSH sliding clamps, though MLH is not absolutely required in vitro for 5' MMR. In other models, MSH remains at the mismatch to authorize excision or can load multiple MLH clamps onto the DNA in the vicinity of the mismatch (not shown). (C) In the sliding clamp model, the EXO1/MSH complex dissociates after excising several hundred nucleotides. Iterative rounds of MSH-EXO1 excision create an excision tract coated with RPA that extends from the 5′ nick to just beyond the mismatch. MLH may limit excision by modulating the number of MSH clamps on DNA. (D) RFC (not shown) loads PCNA clamps with specific orientation at 3′ termini of strand breaks or gaps, and PCNA facilitates high-fidelity DNA synthesis by Pol δ or ε. (E) DNA ligase I seals the nick. The right side of the schematic depicts 3′ MMR. (A) MSH recognizes a mismatch. (B) In the sliding clamp model, ATP-dependent binding and nucleotide switching creates MSH sliding clamps that diffuse from the mismatch. The interaction of ATP-bound MLH heterodimers with MSH sliding clamps and PCNA oriented with respect to 3′ termini activates MLH strand-specific nicking. Alternatively, ATP-activated MSH may remain at the mismatch to load MLH and activate nicking (not shown). (C) Excision is EXO1-dependent or -independent, leading to an RPA-coated excision track. An EXO1-independent Pol δ strand-displacement pathway is not shown. (D) Pol δ or ε with the aid of PCNA completes gap filling. (E) DNA ligase I seals the nick.

[108] **FIGs. 9A-9C** provide a schematic of mismatch repair of PE2 intermediates. MMR inhibition provides additional time for flap ligation, removing the strand discrimination signal for repair of the heteroduplex.

[109] **FIG. 10** shows that expression of dominant-negative MLH1 mutants boosts PE2 efficiency. MLH1 dominant-negative mutants improve PE2 efficiency by 2- to 4-fold. *RNF2* +3 G to C is not responsive to MMR-inhibition.

[110] **FIGs. 11A-11B** show the effect of MLH1 mutants on PE3. MLH1 mutants reduce PE3 indels by half.

[111] **FIGs. 12A-12B** show that MLH1 mutant improvements translate to other sites. **FIG. 12A** shows that PE2 editing efficiency increases with MLH1 mutants, and only *RNF2* +3 G to C is resistant to MMR-inhibition. **FIG. 12B** shows that MLH1 mutants reduce the occurrence of indels by half.

[112] **FIG. 13** provides a schematic showing mismatch repair of PE3 intermediates.

[113] **FIG. 14** provides a schematic showing that mismatch repair differentially resolves PE3 intermediates. Mismatch repair is required for the one edit-favored intermediate.

[114] **FIG. 15A-15H** show screening of MLH1 mutants for smaller size and improved activity. **FIG. 15A** shows that MLH1 Δ754-756 most strongly promotes PE2 editing (hereafter named MLH1dn). MLH1 N-terminal domain approaches the effectiveness of MLH1dn (hereafter named MLH1dn^{NTD}). MLH1 dominant negative mutants may function by saturating binding of MutS. **FIG. 15B** shows that the MLH1 N-terminal domain + NLS approaches the activity of MLH1neg. **FIG. 15C** shows that MLH1dn fusion to PE by a self-cleavable P2A linker (PE-2A-MLH1dn) can improve prime editing efficiency. **FIGs. 15D-15F** show that MMR KD phenocopies MLH1neg expression. **FIGs. 15G-15H** show that the efficiency of PE2 and PE3 is equal in the absence of MMR, suggesting that the complementary nick only serves to bias MMR.

[115] **FIG. 16** shows that MLH1dn reduces indels for PE3. Silent pegRNA is pegRNA that does not encode an edit or produce a mismatch. MLH1dn only reduces PE3 indels if a mismatch is generated.

[116] **FIG. 17** show that mismatch repair of PE heteroduplexes produces a diffuse indel pattern. Indel distribution is broad for PE3 for these edits, but inhibiting MMR with MLH1dn narrows that distribution. This suggests that MMR makes incisions after mismatch recognition that contribute to the indels generated by PE3.

[117] **FIG. 18** shows mismatch repair of PE3 intermediates.

[118] **FIGs. 19A-19B** show that MMR excision of the target locus generates indels in PE3.

[119] **FIGs. 20A-20B** show that MMR knockdown or knockout has no effect on *RNF2* +3 G to C. This suggests that the *RNF2* site is not repaired by MMR or the resulting C:C mismatch is not repaired by MMR.

[120] **FIGs. 21A-21C** show that other substitution edits at *RNF2* can be improved with MLH1dn.

[121] **FIGs. 22A-22B** show that MLH1dn improves substitution edits at other sites, including HEK3. MLH1dn strongly enhances PE2 editing and lowers PE3 indels.

[122] **FIGs. 23A-23D** show that MLH1dn improves substitution edits at other sites, including FANCF. MLH1dn strongly enhances PE2 editing and lowers PE3 indels.

[123] **FIGs. 24A-24B** show that PE improvement by MHL1dn is mismatch dependent. MLH1dn increases PE2 editing by 2-fold on average in HEK293T cells. **FIG. 24A** shows that G to C edits (C:C mismatches) are unaffected by MMR in HEK293T cells. This suggests that G to C edits have a higher baseline efficiency than other substitutions. **FIG. 24B** shows a substantial increase in the ratio of edit:indel purity from MLH1dn used with PE3, which is also mismatch dependent.

[124] **FIGs. 25A-25D** show that MLH1dn also improves the efficiency of small insertion and deletion edits. MMR is known to repair insertions and deletions <15 nucleotides in length.

[125] **FIGs. 26A-26B** show that MLH1dn reduced pegRNA scaffold integration. Scaffold integration events at these sites occur through a double-strand break (DSB) intermediate.

[126] **FIG. 27** shows that MLH1dn does not promote substantial PE off-target editing. Small increases in off-target (OT) editing were observed at the *HEK4* off-target site 3.

[127] **FIGs. 28A-28B** show that MLH1dn does not induce detectable microsatellite instability at biomarker loci. MMR inhibition is known to cause shortening of homopolymer microsatellite regions.

[128] **FIG. 29** shows that MLH1dn offers a method to increase prime editing efficiency at sites without good ngRNAs, such as *HEK4*.

[129] **FIG. 30** shows that MLH1dn improves PE at disease sites.

[130] **FIG. 31** shows that MLH1dn enhances installation of the protective APOE Christchurch allele in mouse astrocytes. A 50% boost in editing efficiency and a large reduction in indels is shown.

[131] **FIG. 32** shows that HEK293T cells are MMR-compromised. The MLH1 promoter is hypermethylated in HEF293T, resulting in lower MLH1 expression.

[132] **FIGs. 33A-33B** show that MLH1dn enhances prime editing in HeLa cells. **FIG. 33A** shows prime editing with PE2. **FIG. 33B** shows prime editing with PE3.

[133] **FIGs. 34A-34B** show that MLH1dn enhances prime editing in HeLa cells. **FIG. 34A** shows editing of *PRNP* +6 G to T. **FIG. 34B** shows editing of APOE +6 G to T and +10 C to A.

[134] **FIGs. 35A-35B** show that MLH1dn has a larger effect in MMR competent cell lines like HeLa.

[135] **FIGs. 36A-36D** show that MLH1dn improvements synergize with stabilized pegRNAs.

[136] **FIGs. 37A-37B** show that contiguous substitutions are useful as another strategy for evading MMR.

[137] **FIG. 38** shows that MMR does not efficiently repair 3 or more contiguous substitutions. Contiguous substitutions therefore offer a method for circumventing MMR and boosting PE efficiency.

[138] **FIGs. 39A-39C** show that MLH1neg improves PE in HeLa cells.

[139] **FIGs. 40A-40G** show that pooled Repair-seq CRISPRi screens reveal genetic determinants of substitution on prime editing outcomes. **FIG. 40A** shows that prime editing with the PE2 system is mediated by the PE2 enzyme (*Streptococcus pyogenes* Cas9 (SpCas9) H840A nickase fused to a reverse transcriptase) and a prime editing guide RNA (pegRNA). The PE3 system uses an additional single guide RNA (sgRNA) to nick the non-edited strand and yield higher editing efficiency. PBS, primer binding site. RT template, reverse transcription template. **FIG. 40B** provides an overview of prime editing Repair-seq CRISPRi screens. A library of CRISPRi sgRNAs and a pre-validated prime edit site are transduced into CRISPRi cell lines and transfected with prime editors targeting the edit site. CRISPRi sgRNA identities and prime edited sites are amplified together from genomic DNA and paired-end sequenced together to link each genetic perturbation with editing outcome. SaCas9, *Staphylococcus aureus* Cas9. **FIG. 40C** shows the effect of each CRISPRi sgRNA on the percentage of sequencing reads reporting the intended G•C-to-C•G prime edit at the targeted edit site in pooled CRISPRi screens. Each value depicts all sequencing reads carrying the same CRISPRi sgRNA. **FIG. 40D** shows the effect of CRISPRi sgRNAs on editing efficiency in all screen conditions. Black dots represent individual non-targeting sgRNAs, black lines show the mean of all non-targeting sgRNAs, and gray

shading represents kernel density estimates of the distributions of all sgRNAs. **FIGs. 40E–40G** show comparisons of gene-level effects of CRISPRi targeting on the intended G•C-to-C•G prime edit across different screen conditions. (**FIG. 40E**) K562 PE2 vs. HeLa PE2. (**FIG. 40F**) K562 PE3+50 vs. HeLa PE3+50. (**FIG. 40G**) K562 PE2 vs. K562 PE3+50. The effect of each gene is calculated as the average log2 fold change in frequency from non-targeting sgRNAs for the two most extreme sgRNAs targeting the gene. Plotted quantities are the mean of n=2 independent biological replicates for each cell type, with bars showing the range of values spanned by the replicates. Black dots represent 20 random sets of three non-targeting sgRNAs.

[140] **FIGs. 41A-41J** show genetic modulators of unintended prime editing outcomes. **FIGs. 41A-41D** show representative examples of four categories of unintended prime editing outcomes observed in CRISPRi screens. In each panel, the black bar depicts the sequence of an editing outcome, the blue bar depicts genomic sequence around the targeted editing site, and the orange bar depicts the pegRNA sequence. Blue and orange lines between the editing outcome and the genome or pegRNA depict local alignments between the outcome sequence and the relevant reference sequence. Mismatches in alignments are marked by X's, and insertions are marked by downward dimples. The location of the programmed edit is marked by a grey box. Red and cyan rectangles on the genome mark SaCas9 protospacers and PAMs, and black vertical lines mark the locations of SaCas9 nick sites. Orange, beige, grey, and red rectangles on the pegRNA mark the primer binding site (PBS), reverse transcription template (RTT), scaffold, and spacer, respectively. **FIGs. 41E-41F** provide a summary of editing outcome categories observed in PE2 screens (**FIG. 41E**) and in PE3+50 screens (**FIG. 41F**) in K562 cells. Plotted quantities are the mean ± SD of all sgRNAs for each indicated gene (60 non-targeting sgRNAs, three sgRNAs per targeted gene), averaged across n=2 independent biological replicates. **FIGs. 41G-41H** show a comparison of the effects of knockdown of all genes targeted in CRISPRi screens on the frequency of joining of reverse transcribed sequence at unintended locations (**FIG. 41G**) or the frequency of deletions (**FIG. 41H**) from PE3+50. The effect of each gene is calculated as the average log2 fold change in frequency from non-targeting sgRNAs for the two most extreme sgRNAs targeting the gene. Plotted quantities are the mean of n=2 independent biological replicates for each cell type, with bars showing the range of values spanned by the replicates. Black dots represent 20 random sets of three non-targeting sgRNAs. **FIG. 41I** shows the frequency of deletion as a function of genomic position relative to programmed PE3+50 nicks

(dashed vertical lines) in K562 screen replicate 1 across all reads for indicated sets of CRISPRi sgRNAs (black line: 60 non-targeting sgRNAs; orange and green lines: three sgRNAs targeting each of MSH2, MSH6, MLH1, and PMS2) (top). Log2 fold change in frequency of deletion as a function of genomic position from MSH2, MSH6, MLH1, and PMS2 sgRNAs compared to non-targeting sgRNAs (bottom). **FIG. 41 J** shows the effect of gene knockdowns on the fraction of all observed deletions that remove sequence at least 25-nt outside of programmed PE3+50 nicks in K562 screens. Each dot represents all reads for all sgRNAs targeting each gene. Black dots represent 20 sets of three random non-targeting sgRNAs.

[141] **FIGs. 42A-42D** show a model for mismatch repair of prime editing intermediates. **FIG. 42A** shows a model for DNA mismatch repair (MMR) of PE2 intermediates. MMR excises and replaces the nicked strand during repair of the prime editor-generated heteroduplex substrate. Infrequent ligation of the nick before MMR recognition deprives the strand discrimination signal for MMR, resulting in un-biased resolution of the heteroduplex. **FIG. 42B** shows a model for MMR of PE3 intermediates. PE3 installs an additional nick on the non-edited strand that can direct MMR to replace the non-edited strand. Ligation of the edited strand nick leaves only the complementary-strand nick to signal repair by MMR, resulting in the desired prime editing outcome. **FIG. 42C** shows prime editing efficiencies of PE2 and PE3 prime editors at endogenous sites (*HEK3*, *EMX1*, and *RUNX1*) in HEK293T cells pre-treated with knockdown siRNAs against MSH2, MSH6, MLH1, or PMS2 transcripts. Cells were pre-transfected with siRNAs 3 days prior to transfection with prime editor components and siRNAs. Genomic DNA was harvested 3 days following transfection with prime editors and additional siRNA, then sequenced. Bars represent the mean of n=3 independent biological replicates. **FIG. 42D** shows prime editing efficiencies in HAP1 ΔMSH2 and HAP1 ΔMLH1 cells (mean of n=3 independent biological replicates). Δ, gene knockout.

[142] **FIGs. 43A-43F** show that engineered dominant negative MMR proteins (dominant negative variants of MSH2, MSH6, PMS2, and MLH1) enhance prime editing. **FIG. 43A** shows editing improvement at *HEK2*, *EMX1*, and *RUNX1* sites by co-expression of PE2 *in trans* with human MMR proteins or dominant negative variants in HEK293T cells. MMR proteins include MSH2, MSH6, PMS2, and MLH1. Dominant negative variants are designated as MSH2 K675R, MSH6 K1140R, PMS2 E41A, PMS2 E705K, MLH1 E34A, and MLH1 Δ756. All values from n = 3 independent biological replicates are shown. **FIG. 43B** shows functional annotation of the

756-aa human MLH1 protein, including an ATPase domain, MSH2 interaction domain, NLS domain, PMS2 dimerization domain, and an endonuclease domain. **FIG. 43C** shows editing enhancement of MLH1 variants co-expressed with PE2 in HEK293T cells at *HEK3*, *EMX1*, and *RUNX1* sites. Red boxes indicate mutations that inactivate MLH1 ATPase or endonuclease function. MLH1dn, MLH1 Δ754-756. MLH1NTD–NLS, codon-optimized MLH1 1-335–NLSSV40. All values from n=3 independent biological replicates are shown. **FIG. 43D** shows a comparison of the top three dominant negative MLH1 variants at additional prime edits. All values from n=3 independent biological replicates are shown. **FIG. 43E** shows prime editing with PE2 and MLH1dn *in trans*, PE2 and $MLH1^{NTD}$–NLS *in trans*, and PE2–P2A–MLH1dn (human codon optimized) in HEK293T cells. Bars represent the mean of n = 3 independent biological replicates. **FIG. 43F** compares the structure of PE2, PE3, PE4, and PE5. In particular, the PE4 editing system consists of a prime editor enzyme (nickase Cas9-RT fusion), MLH1dn, and pegRNA. The PE5 editing system consists of a prime editor enzyme, MLH1dn, pegRNA, and second-strand nicking sgRNA. **FIG. 43G** shows editing efficiencies of PE2, PE3, PE4, and PE5 systems in HEK293T cells. Bars represent the mean of n = 3 independent biological replicates).

[143] **FIGs. 44A-44G** show the characterization of PE4 and PE5 across diverse prime editing classes and cell types. **FIG. 44A** provides a summary of prime editing enhancement by PE4 and PE5 compared to PE2 and PE3 for 84 single-base substitution edits (seven for each substitution type) across seven endogenous sites in HEK293T cells. The grand mean ± SD of all individual values of n = 3 independent biological replicates are shown. **FIG. 44B** shows installation of single base mutations at the FANCF locus with PE2, PE3, PE4, and PE5 in HEK293T cells. Bars represent the mean of n = 3 independent biological replicates. **FIG. 44C** shows that PE4 improves the 1- and 3-bp insertion and deletion prime edits compared to PE2 in HEK293T cells. Bars represent the mean of n = 3 independent biological replicates. **FIG. 44D** shows PE4 editing enhancement over PE2 across 33 different insertion and deletion prime edits. Bars represent the mean of all individual values of n=3 independent biological replicates. **FIGs. 44E–44F** provide a summary of PE2 and PE4 editing efficiencies for 35 different substitutions of 1 to 5 contiguous bases at five endogenous sites in HEK293T cells. Seven pegRNAs were tested for each number of contiguous bases altered. The mean ± SD of all individual values of n = 3 independent biological replicates are shown. **FIGs. 44G-44H** show that installation of additional silent or

benign mutations near the intended edit can increase editing efficiency by generating a heteroduplex substrate that evades MMR. The PAM sequence (NGG) for each target is underlined. The amino acid sequence of the targeted gene is centered above each DNA codon. Values represent the mean±SD of n=3 independent biological replicates. **FIG. 44I** shows a comparison of prime editing enhancement in different cell types. PE4 and PE5 systems enhance prime editing to a greater extent in MMR deficient cells (MMR–) than in MMR proficient cells (MMR+). The same set of 30 pegRNAs encoding single-base substitution edits were tested in HEK293T and HeLa cells. K562 and U2OS cells were edited with 10 pegRNAs that are a direct subset of the 30 pegRNAs tested in HEK293T and HeLa cells. The mean ± SD of all individual values of sets of n = 3 independent biological replicates are shown. P values were calculated using the Mann-Whitney U test. **FIG. 44J** shows prime editing with PE2, PE3, PE4, and PE5 in HeLa, K562, and U2OS cells. Bars represent the mean of n = 3 independent biological replicates).

[144] FIGs. 45A-45H show the effect of dominant negative MLH1 on prime editing product purity and off-targeting. **FIG. 45A** shows that edit-encoding pegRNAs program a base change within the nascent 3′ DNA flap and generate a heteroduplex following flap interconversion. Non-editing pegRNAs template a 3′ DNA flap with perfect complementarity to the genomic target site. **FIG. 45B** shows the frequency of indels from PE3 or PE5 with four edit-encoding pegRNAs that program single base mutations or four non-editing pegRNAs. Short horizontal bars indicate the mean of all individual values of sets of n = 3 independent biological replicates. **FIG. 45C** shows the ratio of indel frequency from PE5 over PE3 with 4 edit-encoding pegRNAs that program single base mutations or four edit-encoding pegRNAs that program single base mutations or four non-editing pegRNAs. Short horizontal bars indicate the mean of all individual values of sets of n = 3 independent biological replicates. **FIG. 45D** shows distribution of deletions at genomic target DNA formed by PE3 and PE5 using 12 substitution-encoding pegRNAs at endogenous DNMT1 and RNF2 loci in HEK293T cells. Dotted lines indicate position of pegRNA- and sgRNA-directed nicks. Data represent the mean ± SD of n=3 independent biological replicates. **FIG. 45E** shows PE5/PE3 ratio of frequency of deletions that remove sequence greater than 25-nt outside of pegRNA- and sgRNA-directed nicks in HEK293T cells. Each dot represents one of 84 total pegRNAs that program substitution edits at a combined seven loci (mean of n=3 independent biological replicates). **FIG. 45F** shows PE5/PE3 ratio of

frequency of editing outcomes with unintended pegRNA scaffold sequence incorporation or unintended flap rejoining in HEK293T cells. Each dot represents one of 84 total pegRNAs that program substitution edits at a combined seven loci (mean of n = 3 independent biological replicates). **FIG. 45G** shows off-target prime editing by PE2 and PE4 in HEK293T cells. Bars represent the mean of n = 3 independent biological replicates). **FIG. 45H** shows high-throughput sequencing analysis of 17 sensitive microsatellite repeat loci used for clinical diagnosis of MMR deficiency. HAP1 and HeLa cells are MMR-proficient, and HCT116 cells have impaired MMR. HAP1 ΔMSH2 cells underwent 60 cell divisions following MSH2 knockout. HeLa cells were transiently transfected with PE2 or PE4 components and incubated for 3 days before sequencing. wt, wild-type. All values from n=2 independent biological replicates are shown.

[145] **FIGs. 46A-46F** show that PEmax architecture with PE4 and PE5 editing systems enhances editing at disease-relevant gene targets and cell types. **FIG 46A** shows a schematic of PE2 and PEmax editor architectures. bpNLSSV40, bipartite SV40 NLS nuclear localization signal. MMLV RT, Moloney Murine Leukemia Virus reverse transcriptase pentamutant; codon opt., human codon-optimized. **FIG. 46B** shows that engineered pegRNAs (epegRNAs) contain a 3′ RNA structural motif that improve prime editing performance. **FIG. 46C** shows prime editing efficiencies of PE4 and PE5 combined with PEmax architectures and epegRNAs. Seven single-base substitution edits targeting different loci were tested in HeLa and HEK293T cells. Fold changes indicate the average of fold increases from each edit tested. The mean±SD of all individual values of n=3 independent biological replicates are shown. **FIG. 46D** shows prime editing at therapeutically-relevant sites in wild-type HeLa and HEK293T cells. The HBB locus is edited at the E6 codon commonly mutated in patients with sickle cell disease (E6V). The CDKL5 edit is at a site for which the c.1412delA mutation causes CDKL5 deficiency disorder. epegRNAs were used for editing the HBB, PRNP, and CDKL5 loci. Bars represent the mean of n=3 independent biological replicates. **FIG. 46E** shows correction of CDKL5 c.1412delA via an A•T insertion and a silent G•C-to-A•T edit in iPSCs derived from a patient heterozygous for the allele. Editing efficiencies indicate the percentage of sequencing reads with c.1412delA correction out of editable alleles that carry the mutation. Indel frequencies reflect all sequencing reads that contain any indels. Bars represent the mean of n=3 independent biological replicates. **FIG. 46F** shows prime editing in primary human T cells. Bars represent the mean of n=3 independent biological replicates from different healthy T cell donors.

[146] **FIGs. 47A-47J** show the design and results of Repair-seq screens for substitution prime editing outcomes. **FIG. 47A** shows optimization of a *Staphylococcus aureus* (Sa)-pegRNA for installation of a G•C-to-C•G edit within a lentivirally integrated HBB sequence using SaPE2 in HEK293T cells. PBS, primer-binding site. Data represent the mean of n = 3 independent biological replicates. **FIG. 47B** shows the design of the prime editing Repair-seq lentiviral vector (pPC1000). In Repair-seq screens, a 453-bp region containing CRISPRi sgRNA sequence and prime editing outcome is amplified from genomic DNA for paired-end Illumina sequencing. The CRISPRi sgRNA is sequenced with a 44-nt Illumina forward read (R1), and the prime edited site (including +50 and –50 nick sites) is sequenced with a 263-nt Illumina reverse read (R2). Black triangles indicate positions of SaPE2-induced nicks programmed by Sa-pegRNA and Sa-sgRNAs. Sizes of all vector components are to scale. **FIG. 47C** shows a schematic of PE2, PE3+50, and PE3–50 prime editing configurations with SaPE2 protein (SaCas9 N580A fused to an engineered MMLV RT). **FIG. 47D** shows validation of intended G•C-to-C•G editing at the lentivirally-integrated Repair-seq edit site in HeLa cells expressing dCas9–BFP–KRAB cells. Bars represent the mean of n = 2 independent biological replicates. **FIG. 47E** shows prime editing at the Repair-seq edit site with and without blasticidin selection in HeLa cells expressing dCas9–BFP–KRAB. SaPE2–P2A–BlastR prime editor was used for all conditions. Bars represent the mean of n = 2. **FIG. 47F** shows functional annotation classes of the genes targeted by the pooled CRISPRi sgRNA library used in Repair-seq screens. **FIGs. 47G-47J** show that the knockdown of MSH2, MSH6, MLH1, and PMS2 increases the frequency of the intended +6 G•C-to-C•G prime edit in all Repair-seq screens. Dots represent reads from individual CRISPRi sgRNAs.

[147] **FIGs. 48A-48I** show the genetic modulators of unintended prime editing outcomes. **FIG. 48A** shows an overview of PE3–50 outcomes in HeLa CRISPRi screens. TP53BP1 knockdown dramatically reduces formation of all unintended editing outcomes. **FIG. 48B** shows additional details of PE2 outcomes in K562 CRISPRi screens, supplementing **FIG. 41H. FIG. 48C** shows additional details of PE3+50 outcomes in K562 CRISPRi screens, supplementing information in FIG. 41G. **FIGs. 48D-48I** show comparisons of effects of gene knockdown on frequencies of indicated outcome categories in indicated screen conditions. Plotted quantities are the mean of the log2 fold changes from non-targeting sgRNAs for the two most extreme sgRNAs per gene, averaged over n = 2 independent biological replicates per condition. Error bars mark the range of

values spanned by the replicates. Black dots represent 20 random sets of three non-targeting sgRNAs. **FIG. 48D** shows that MSH2, MLH1, and PMS2 knockdown produce larger fold changes in installation of additional edits than in intended edits in K562 PE2 screens. **FIG. 48E** shows unintended joining of reverse transcribed sequence in PE2 screens in K562 and HeLa cells are most increased by knockdown of Fanconi anemia genes (red) as well as a set of RAD51 homologs and other genes involved in homologous recombination (blue). **FIG. 48F** shows deletions in in PE2 screens in K562 and HeLa cells are most increased by a set of RAD51 homologs and other genes involved in homologous recombination (blue). **FIG. 48G** shows that in addition to MSH2, MLH1, and PMS2, HLTF knockdown produces larger fold changes in installation of additional edits than in intended edits in K562 PE3+50 screens. **FIG. 48H** shows that tandem duplications in HeLa and K562 PE3+50 screens are most decreased by knockdown of POLD and RFC subunits. **FIG. 48I** shows deletions in HeLa PE3+50 and PE3–50 screens have dramatically divergent genetic regulators, highlighting differences in the processing of the different overhang configurations.

[148] **FIGs. 49A-49F** show validation of prime editing Repair-seq screen results. **FIGs. 49A-49B** show alignment of Sa-pegRNAs, their templated 3′ DNA flaps following SaPE2 reverse transcription, and the genomic target sequence (top). Compared to the Sa-pegRNA used in Repair-seq screens (**FIG. 49A**), an Sa-pegRNA with recoded scaffold sequence (**FIG. 49B**) templates an extended 3′ DNA flap with reduced homology with genomic target sequence. The recoded Sa-pegRNA contains 2 base pair changes that preserve base pairing interactions within the scaffold. Reverse transcription of the Sa-pegRNA scaffold can generate a misextended 3′ flap that is incorporated into the genome. Vertical lines depict base pairing. X's depict mismatches between the misextended reverse-transcribed 3′ flap and genomic sequence. **FIGs. 49A-49B** also show frequencies of editing outcome categories observed at the screen edit site from arrayed PE2 and PE3+50 experiments in HeLa CRISPRi cells (bottom). Prime editing with the Sa-pegRNA used in siteRepair-seq screens (**FIG. 49A**) or a recoded Sa-pegRNA (**FIG. 49B**) results in different frequencies of installation of unintended edits from nearly-matched scaffold. Plotted quantities are the mean±SD of n=4 independent biological replicates, for each cell line containing MSH2 or non-targeting CRISPRi sgRNAs. **FIG. 49C** shows the mechanism of DNA mismatch repair in humans. **FIG. 49D** shows mismatch repair of a prime editing heteroduplex intermediate could install additional non-programmed nicks from MutLα endonuclease activity.

Excision from these non-programmed nicks and subsequent repair of the resulting intermediates may contribute to larger and more frequent indel byproducts observed from MMR activity. **FIG. 49E** shows the knockdown efficiency of siRNA treatment relative to a non-targeting siRNA control in HEK293T cells. Cells were transfected with siRNAs, incubated for 3 days, transfected with PE2, pegRNAs, and the same siRNAs, then incubated for another 3 days before relative RNA abundances were assayed by RT-qPCR. NT, non-targeting. Data represent the mean of n=3 independent biological replicates. Each dot represents the mean of n = 3 technical replicates. **FIG. 49F** shows editing in HEK293T cells co-transfected with prime editor components and siRNAs. Cells were not pre-treated with siRNAs before transfection with prime editor. Bars represent the mean of n = 3 independent biological replicates.

[149] FIGs. 50A-50H show the development and characterization of dominant negative MMR proteins that enhance prime editing. **FIG. 50A** shows the prime editing efficiencies from MMR proteins or dominant negative variants expressed in trans with or fused directly to PE2 in HEK293T cells. 32aa linker, (SGGS)×2–XTEN–(SGGS)×2 (SEQ ID NO: 125) (SGGSSGGSSGSETPGTSESATPES SGGSSGGS (SEQ ID NO: 125) or structurally, [SGGS]-[SGGS]-[SGSETPGTSESATPES]-[SGGS]-[SGGS] (SEQ ID NO: 125)). codon opt., human codon optimized. Data within the same graph originate from experiments performed at the same time. Data represent the mean ± SD of n = 3 independent biological replicates. **FIG. 50B** shows titration of MLH1dn plasmid and PE2 plasmid transfection doses in HEK293T cells. Maximum plasmid amounts tested were 200 ng PE2 and 100 ng MLH1dn. Data represent the mean ± SD of n = 3 independent biological replicates. **FIG. 50C** shows prime editing with MLH1dn co-expression in MMR-deficient HCT116 cells that contain a biallelic deletion in MLH1. Bars represent the mean of 3 replicates. **FIG. 50D** shows a comparison of prime editing with human MLH1dn (human codon-optimized) or mouse MLH1dn (mouse codon optimized) in human HEK293T cells. Bars represent the mean of n = 3 independent biological replicates. **FIG. 50E** shows a comparison of prime editing with human MLH1dn (human codon optimized) or mouse MLH1dn (mouse codon optimized) in mouse N2A cells. Bars represent the mean of n = 3 replicates. **FIG. 50F** shows that *MLH1* knockout in clonal HeLa cell lines enhances prime editing efficiency to a greater extent than MLH1 co-expression in clonal wild-type HeLa cells. Δ, knockout. Bars represent the mean of n = 3 or 4 independent biological replicates. **FIG. 50G** shows editing at the FANCF locus with PE3b and PE5b (complementary-strand nick that is

specific for the edited sequence) in HEK293T cells. PE5b, PE3b editing system with MLH1dn co-expression. Bars represent the mean of n = 3 independent biological replicates. **FIG. 50H** shows editing at the HEK2 locus with complementary-strand nicks in HEK293T cells. "None" indicates the lack of a nick, which denotes a PE2 or PE4 editing strategy. Bars represent the mean of n = 3 independent biological replicates.

[150] **FIGs. 51A-51J** show the characterization of PE4 and PE5 across diverse prime edit classes and cell types. **FIG. 51A** shows a comparison of PE2, PE3, PE4, and PE5 for 84 single-base substitution prime edits across seven endogenous sites in HEK293T cells. Bars represent the mean of n = 3 independent biological replicates). **FIG. 51B** provides a summary of PE4 enhancement in editing efficiency over PE2 for 84 single-base substitution edits across seven endogenous sites in HEK293T cells. PE4/PE2 fold improvements may be lower for PAM edits due to the high basal editing efficiency for PAM edits or the high representation of G•C-to-C•G edits (five out of 15 in this category). Data represent the mean ± SD of n = 3 independent biological replicates. **FIG. 51C** shows the efficiencies of single-base substitution prime edits that alter the PAM (+5 G and +6 G bases) of prime editing target protospacers in HEK293T cells. Four G•C-to-A•T, five G•C-to-C•G, and six G•C-to-T•A PAM edits across a combined seven endogenous sites are shown. The mean of all individual values of n = 3 independent biological replicates are shown. **FIG. 51D** shows the effect of siRNA knockdown of MMR genes on G•C-to-C•G editing at the RNF2 locus in HEK293T cells. Bars represent the mean of n = 3 independent biological replicates. **FIG. 51E** shows the effect of MMR gene knockout on G•C-to-C•G editing at the RNF2 locus in HAP1 cells. Δ, gene knockout. Bars represent the mean of n = 3 independent biological replicates. **FIG. 51F** shows prime editing at the integrated screen edit site with CRISPRi knockdown in HeLa CRISPRi cells. PE2 indicates editing with SaPE2 protein and Sa-pegRNA. PE3+50 indicates editing with SaPE2 protein, Sa-pegRNA, and Sa-sgRNA that programs a +50 complementary-strand nick. Bars represent the mean of n = 5 independent biological replicates. **FIG. 51G** provides a summary of PE5 enhancement in editing efficiency over PE3 for 84 single-base substitution edits in HEK293T cells. The grand mean ± SD of all individual values of n = 3 independent biological replicates are shown. **FIG. 51H** shows PE4 enhancement in editing efficiency over PE2 across a range of insertion and deletion prime edit lengths in HEK293T cells. A total of 33 different prime edits at a combined three endogenous loci are shown. The mean of all individual values of n = 3 independent biological replicates are

shown. **FIG. 51I** shows that PE5 improves editing efficiency and reduces indel byproducts compared to PE3 across small insertion and deletion prime edits in HEK293T cells. **FIG. 51J** shows PE2 and PE4 editing efficiencies at 33 different insertion and deletion prime edits across a combined three endogenous loci. Bars represent the mean of all individual values of n=3 independent biological replicates.

[151] **FIGs. 52A-52C** show characterization of PE4 and PE5 systems and improved prime editing efficiency with additional silent mutations. **FIG. 52A** shows substitutions of contiguous bases with PE2 and PE4 in HEK293T cells. The top sequence indicates the original, unedited genomic sequence. Numbers denote the position of the edited nucleotide relative to the PE2 nick site. Nucleotides within the SpCas9 PAM sequence (NGG) are underlined. Sequences of the intended edited product are shown below, with edited nucleotides marked in red. Bars represent the mean of n = 3 independent biological replicates. **FIG. 52B** shows that installation of additional silent mutations can increase prime editing efficiency by evading MMR. PE4/PE2 fold-change in editing frequency reflects the extent to which MMR activity impedes the indicated prime edit. Edited nucleotides that make the indicated coding mutation are marked in red, and edited nucleotides that make silent mutations are marked in green. Data represent the mean±SD of n=3 independent biological replicates. **FIG. 52C** shows installation of 22 single-base substitution prime edits across seven endogenous sites in HeLa cells with PE2, PE3, PE4, and PE5. Bars represent the mean of n=3 independent biological replicates.

[152] **FIGs. 53A-53G** show the effect of dominant negative MLH1 on prime editing product purity and off-targeting. **FIG. 53A** shows the frequency of indels in HEK293T cells treated with pegRNAs, nicking sgRNAs, and PE2 enzyme, RT-impaired PE2 (PE2–dRT), or nickase Cas9 (SpCas9 H840A), with and without MLH1dn. Non-editing pegRNAs encode a 3′ DNA flap with perfect homology to the genomic target. Bars represent the mean of n = 3 independent biological replicates. **FIG. 53B** shows the distribution of deletion outcomes from PE3 and PE5 at endogenous loci in HEK293T cells. 12 different pegRNAs that program single-base substitutions were tested at each indicated locus. Dotted lines indicate position of pegRNA- and sgRNA-directed nicks. Data represent the mean ± SD of n = 3 independent biological replicates. **FIG. 53C** shows the distribution of deletion outcomes from PE3 and PE5 with an edit-encoding and non-editing pegRNA in HEK293T cells. The non-editing pegRNA templates a 3′ DNA flap with perfect complementarity to the genomic target sequence. Data represent the mean ± SD of n = 3

independent biological replicates. **FIG. 53D** shows the frequency of all prime editing outcomes with unintended pegRNA scaffold sequence incorporation or unintended flap rejoining in HEK293T cells. 12 pegRNAs each programming a different single-base substitution were tested at each of the seven indicated loci. Each dot represents an individual pegRNA at the indicated loci (mean of n = 3 independent biological replicates). **FIG. 53E** shows the off-target prime editing by PE2 and PE4 in HEK293T cells (mean of n = 3 independent biological replicates). **FIG. 53F** shows the distribution and cumulative distribution of microsatellite repeat lengths in the indicated cell types and treatments. HAP1 and HeLa cells are MMR-proficient, and HCT116 cells have impaired MMR. HAP1 ΔMSH2 cells underwent 60 cell divisions following knockout of MSH2. HeLa cells were transiently transfected with PE2 or PE4 components and grown for a following 3 days before sequencing. wt, wild-type. All values from n = 2 independent biological replicates are shown. **FIG. 53G** shows prime editing at the on-target locus in HeLa cells transfected with PE2 or PE4 components. Bars represent the mean of n = 2 independent biological replicates. Microsatellite lengths were assayed from genomic DNA taken from these PE2 and PE4-treated HeLa cells.

[153] **FIGs. 54A-54F** show that use of PEmax architecture with PE4 and PE5 editing systems enhances editing at disease-relevant gene targets and cell types. **FIG. 54A** shows a schematic of PE2 and PEmax editor architectures. $bpNLS^{SV40}$, bipartite SV40 NLS. MMLV RT, Moloney Murine Leukemia Virus reverse transcriptase pentamutant. GS codon, Genscript human codon optimized. **FIG. 54B** shows engineered pegRNAs (epegRNAs) containing a 3′ RNA structural motif that improve prime editing performance. **FIG. 54C** shows prime editing efficiencies of PE4 and PE5 combined with PEmax architectures and epegRNAs. Seven single-base substitution edits targeting different loci were tested in HeLa and HEK293T cells. Fold changes indicate the average of fold increases from each edit tested. The mean ± SD of all individual values of n = 3 independent biological replicates are shown. **FIG. 54D** shows prime editing at therapeutically-relevant sites in wild-type HeLa and HEK293T cells. The HBB locus is edited at the E6 codon commonly mutated in patients with sickle cell disease (E6V). The CDKL5 edit is at a site for which the c.1412delA mutation causes CDKL5 deficiency disorder. epegRNAs were used for editing the HBB, PRNP, and CDKL5 loci. Bars represent the mean of n = 3 independent biological replicates. **FIG. 54E** shows the correction of *CDKL5* c.1412delA via an A•T insertion and a silent G•C-to-A•T edit in iPSCs derived from a patient heterozygous for the allele. Editing

efficiencies indicate the percentage of sequencing reads with c.1412delA correction out of editable alleles that carry the mutation. Indel frequencies reflect all sequencing reads that contain any indels. Bars represent the mean of n = 3 independent biological replicates. **FIG. 54F** shows prime editing in primary T cells. Bars represent the mean of n = 3 independent biological replicates from different healthy T cell donors.

[154] **FIGs. 55A-55B** show the development of PEmax and application of PE4 and PE5 to primary cell types. **FIGs. 55A-55B** show screening of prime editor variants to maximize editing efficiency in HeLa cells. All prime editor architectures carry a Cas9 H840A mutation to prevent nicking of the complementary DNA strand at the target protospacer. *NLSSV40 contains a 1-aa deletion outside the PKKKRKV (SEQ ID NO: 132) NLSSV40 consensus sequence. All individual values of n = 3 independent biological replicates are shown.

[155] **FIGs. 56A-56G** show development of PEmax and application of PE4 and PE5 to primary cell types. **FIG. 56A** shows a screen of prime editor variants for improved editing efficiency with the PE3 system in HeLa cells. All prime editor architectures carry a SpCas9 H840A mutation to prevent nicking of the complementary DNA strand at the target protospacer. NLSSV40 indicates the bipartite SV40 NLS. *NLSSV40 contains a 1-aa deletion outside the PKKKRKV (SEQ ID NO: 132) NLSSV40 consensus sequence. All individual values of n=3 independent biological replicates are shown. **FIG. 56B** shows the architecture of the original PE2 editor (Anzalone et al., 2019), PE2* (Liu et al., 2021), CMP–PE–V1 (Park et al., 2021), and prime editor variants developed in this work (PEmax, CMP–PEmax). HN1, HMGN1; H1G, histone H1 central globular domain; codon opt., human codon optimized. **FIG. 56C** shows that PEmax outperforms other prime editor architectures with the PE3 system in HeLa cells. Bars represent the mean of n=3 independent biological replicates. **FIG. 56D** shows fold-change in editing efficiency of prime editor architectures compared to PE2 with the PE3 system in HeLa cells. The mean±SD of all individual values of n=3 independent biological replicates are shown. **FIG. 56E** shows the intended editing and indel frequencies from PE4, PE4max (PE4 editing system with PEmax architecture), PE5, and PE5max (PE5 editing system with PEmax architecture) in HeLa and HEK293T cells. Seven substitution prime edits targeting different endogenous loci were tested for each condition. The mean ± SD of all individual values of n = 3 independent biological replicates are shown. **FIG. 56F** shows the correction of CDKL5 c.1412delA via an A•T insertion and a G•C-to-A•T edit in iPSCs derived from a patient

heterozygous for the disease allele. Editing efficiencies indicate the percentage of sequencing reads with c.1412delA correction out of editable alleles that carry the mutation. Indel frequencies reflect all sequencing reads that contain any indels that do not map to the c.1412delA allele or wild-type sequence. 1 μg of PE2 mRNA was used in all conditions shown. Bars represent the mean of n = 3 independent biological replicates. **FIG. 56G** shows prime editing in primary T cells. Bars represent the mean of n = 3 independent replicates from different T cell donors.

[156] **FIGs. 57A-57B** show that the recoded pegRNA scaffold reduces unintended outcomes from scaffold sequence incorporation. **FIG. 57A** shows an alignment of the prime editing Repair-seq target site and SaPE2-generated 3′ DNA flaps templated by (top) the Sa-pegRNA used in Repair-seq screens, or (bottom) an Sa-pegRNA with a recoded scaffold sequence. 3′ flap sequences are aligned with the templated region of the Sa-pegRNA shown above (RT template or scaffold). Red indicates position of the intended +6 G•C to C•G edit programmed by both Sa-pegRNAs. Blue indicates positions at which the genomic target sequence does not align with the 3′ flap sequence templated by the Sa-pegRNA scaffold. Unintended edits from incorporation of a 3′ flap containing a reverse transcribed Sa-pegRNA scaffold sequence may occur at these blue-indicated nucleotides. **FIG. 57B** shows a summary of editing outcome categories observed in PE2 and PE3+50 experiments in HeLa CRISPRi cells. Screen pegRNA indicates the Sa-pegRNA used in prime editing Repair-seq screens. Sa-pegRNA with recoded scaffold (sequence shown in FIG. 54A) avoids sequence homology with the Repair-seq edit site. Plotted quantities are the mean ± SD of one CRISPRi sgRNA for each indicated target (MSH2 and non-targeting), averaged across n = 4 independent biological replicates.

[157] **FIG. 58** shows a comparison of PE3max (PE3 editing system with PEmax protein) and PE3 (PE3 editing system with PE2 protein) in HeLa cells (mean of n = 3 independent biological replicates).

[158] **FIG. 59** shows that PE improvement with MLH1dn depends on prime edit size. MMR most efficiently repairs substitutions and insertion and deletion errors of fewer than or equal to approximately 13 bp in length.

[159] **FIG. 60** shows that PE4 and epegRNAs enable prime editing with a single pegRNA integrant.

[160] **FIG. 61** shows that PE5 improves installation of the protective Christchurch allele in an *APOE4* mouse astrocyte model.

[161] **FIGs. 62A-62C** show that inhibiting p53 enhances the efficiency and precision of PE3 prime editing. This is particularly true when the nicking sgRNA makes a nick upstream (- side) of the pegRNA-directed nick. Each point on the graphs represents an individual CRISPRi gene knockdown in the Repair-seq screens. The axes depict log2 fold changes compared to control. Knocking down TP53BP1 (p53 gene) increases intended editing (x-axis) and decreases three types of unintended editing (y-axes), including joining of the reverse transcribed sequence at unintended locations (FIG. 62A), unintended deletions (FIG. 62B), and unintended tandem duplications (FIG. 62C).

[162] **FIG. 63** shows that a p53 inhibitor (i53) can enhance the efficiency and precision of PE3 prime editing. This is particularly true when the nicking sgRNA makes a nick upstream (- side) of the pegRNA-directed nick. Only the *EMX1* site uses a nick on the "-" side.**FIG. 64** represents various aspects of the disclosure, including the use of CRISPRi screens to reveal cellular genes—including mismatch repair genes—having an impact on prime editing outcomes, the use of engineered MLH1 of the mismatch repair (MMR) pathway to enhance the efficiency and precision of prime editing, and the demonstration that improved prime editing systems described herein (e.g., PE4 and PE5 systems, and PEmax editor) were shown to exhibit the same beneficial effects in many cell types.

[163] **FIG. 64** shows that CRISPRi screens reveal cellular determinants of prime genome editing, that engineered MLH1 protein enhances prime editing efficiency and precision, and that improved prime editing systems were characterized across edit and cell types.

[164] **FIG. 65** provides a schematic showing the optimization of PE2 protein.

[165] **FIG. 66** shows the fold change in the frequency of the intended edit using PE2 and various other PE constructs in HEK293T cells (low plasmid dose) at a range of gene targets (*HEK3, EMX1, RNF2, FANCF, FUNX1, DNMT1, VEGFA, HEK4, PRNP, APOE, CXCR4, HEK3*).

[166] **FIG. 67** shows the fold change in the frequency of the intended edit using PE3 and various prime editor constructs in HeLa cells at a range of gene targets (*HEK3, FANCF, RUNX1, VEGFA*).

[167] **FIG. 68** shows a comparison of prime editing in HEK293T vs. HeLa editing using various PE constructs.

[168] **FIG. 69** shows NLS architecture optimization of PE3 in HeLa cells.

[169] **FIG. 70** provides a schematic showing the final PEmax construct, which corresponds to SEQ ID NO: 99.

[170] **FIG. 71** shows that PEmax increases indels in addition to the intended edit.

DEFINITIONS

[171] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton *et al.*, Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger *et al.* (eds.), Springer Verlag (1991); and Hale & Marham, The Harper Collins Dictionary of Biology (1991). As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

Cas9

[172] The term "Cas9" or "Cas9 nuclease" refers to an RNA-guided nuclease comprising a Cas9 domain, or a fragment thereof (*e.g.*, a protein comprising an active or inactive DNA cleavage domain of Cas9, and/or the gRNA binding domain of Cas9). A "Cas9 domain" as used herein, is a protein fragment comprising an active or inactive cleavage domain of Cas9 and/or the gRNA binding domain of Cas9. A "Cas9 protein" is a full length Cas9 protein. A Cas9 nuclease is also referred to sometimes as a casn1 nuclease or a CRISPR (Clustered Regularly Interspaced Short Palindromic Repeat)-associated nuclease. CRISPR is an adaptive immune system that provides protection against mobile genetic elements (viruses, transposable elements, and conjugative plasmids). CRISPR clusters contain spacers, sequences complementary to antecedent mobile elements, and target invading nucleic acids. CRISPR clusters are transcribed and processed into CRISPR RNA (crRNA). In type II CRISPR systems, correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (rnc), and a Cas9 domain. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves a linear or circular dsDNA target complementary to the spacer. The target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3′-5′ exonucleolytically. In nature, DNA-binding and

cleavage typically requires protein and both RNAs. However, single guide RNAs ("sgRNA", or simply "gNRA") can be engineered to incorporate aspects of both the crRNA and tracrRNA into a single RNA species. See, *e.g.*, Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of which are hereby incorporated by reference. Cas9 recognizes a short motif in the CRISPR repeat sequences (the PAM or protospacer adjacent motif) to help distinguish self versus non-self. Cas9 nuclease sequences and structures are well known to those of skill in the art (see, *e.g.*, "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*." Ferretti *et al.*, J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011); and "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of each of which are incorporated herein by reference). Cas9 orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the organisms and loci disclosed in Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) *RNA Biology* 10:5, 726-737; the entire contents of which are incorporated herein by reference. In some embodiments, a Cas9 nuclease comprises one or more mutations that partially impair or inactivate the DNA cleavage domain.

[173] A nuclease-inactivated Cas9 domain may interchangeably be referred to as a "dCas9" protein (for nuclease-"dead" Cas9). Methods for generating a Cas9 domain (or a fragment thereof) having an inactive DNA cleavage domain are known (see, *e.g.*, Jinek *et al.*, *Science*. 337:816-821(2012); Qi *et al.*, "Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression" (2013) *Cell*. 28;152(5):1173-83, the entire contents of each of which are incorporated herein by reference). For example, the DNA cleavage domain of Cas9 is known to include two subdomains, the HNH nuclease subdomain and the RuvC1

subdomain. The HNH subdomain cleaves the strand complementary to the gRNA, whereas the RuvC1 subdomain cleaves the non-complementary strand. Mutations within these subdomains can silence the nuclease activity of Cas9. For example, the mutations D10A and H840A completely inactivate the nuclease activity of *S. pyogenes* Cas9 (Jinek *et al.*, *Science*. 337:816-821(2012); Qi *et al.*, *Cell*. 28;152(5):1173-83 (2013)). In some embodiments, proteins comprising fragments of Cas9 are provided. For example, in some embodiments, a protein comprises one of two Cas9 domains: (1) the gRNA binding domain of Cas9; or (2) the DNA cleavage domain of Cas9. In some embodiments, proteins comprising Cas9 or fragments thereof are referred to as "Cas9 variants." A Cas9 variant shares homology to Cas9, or a fragment thereof. For example, a Cas9 variant is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, at least about 99.8% identical, or at least about 99.9% identical to wild type Cas9 (e.g., SpCas9 of SEQ ID NO: 2). In some embodiments, the Cas9 variant may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or more amino acid changes compared to wild type Cas9 (e.g., SpCas9 of SEQ ID NO: 2). In some embodiments, the Cas9 variant comprises a fragment of SEQ ID NO: 2 Cas9 (*e.g.*, a gRNA binding domain or a DNA-cleavage domain), such that the fragment is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to the corresponding fragment of wild type Cas9 (e.g., SpCas9 of SEQ ID NO: 2). In some embodiments, the fragment is at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identical, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% of the amino acid length of a corresponding wild type Cas9 (e.g., SpCas9 of SEQ ID NO: 2).

Circular permutant

[174] As used herein, the term "circular permutant" refers to a protein or polypeptide (e.g., a Cas9) comprising a circular permutation, which is a change in the protein's structural configuration involving a change in the order of amino acids appearing in the protein's amino

acid sequence. In other words, circular permutants are proteins that have altered N- and C-termini as compared to a wild-type counterpart, e.g., the wild-type C-terminal half of a protein becomes the new N-terminal half. Circular permutation (or CP) is essentially the topological rearrangement of a protein's primary sequence, connecting its N- and C-terminus, often with a peptide linker, while concurrently splitting its sequence at a different position to create new, adjacent N- and C-termini. The result is a protein structure with different connectivity, but which often can have the same overall similar three-dimensional (3D) shape, and possibly include improved or altered characteristics, including reduced proteolytic susceptibility, improved catalytic activity, altered substrate or ligand binding, and/or improved thermostability. Circular permutant proteins can occur in nature (e.g., concanavalin A and lectin). In addition, circular permutation can occur as a result of posttranslational modifications or may be engineered using recombinant techniques.

Circularly permuted Cas9

[175] The term "circularly permuted Cas9" refers to any Cas9 protein, or variant thereof, that occurs as a circular permutant, whereby its N- and C-termini have been topically rearranged. Such circularly permuted Cas9 proteins ("CP-Cas9"), or variants thereof, retain the ability to bind DNA when complexed with a guide RNA (gRNA). See, Oakes *et al.*, "Protein Engineering of Cas9 for enhanced function," *Methods Enzymol*, 2014, 546: 491–511 and Oakes *et al.*, "CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification," *Cell*, January 10, 2019, 176: 254-267, each of which are incorporated herein by reference. The instant disclosure contemplates any previously known CP-Cas9 or use of a new CP-Cas9 so long as the resulting circularly permuted protein retains the ability to bind DNA when complexed with a guide RNA (gRNA). Exemplary CP-Cas9 proteins are SEQ ID NOs: 54-63.

CRISPR

[176] CRISPR is a family of DNA sequences (i.e., CRISPR clusters) in bacteria and archaea that represent snippets of prior infections by a virus that have invaded the prokaryote. The snippets of DNA are used by the prokaryotic cell to detect and destroy DNA from subsequent attacks by similar viruses and effectively compose, along with an array of CRISPR-associated proteins (including Cas9 and homologs thereof) and CRISPR-associated RNA, a prokaryotic immune defense system. In nature, CRISPR clusters are transcribed and processed into CRISPR

RNA (crRNA). In certain types of CRISPR systems (e.g., type II CRISPR systems), correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (rnc) and a Cas9 protein. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves a linear or circular dsDNA target complementary to the RNA. Specifically, the target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3′-5′ exonucleolytically. In nature, DNA-binding and cleavage typically requires protein and both RNAs. However, single guide RNAs ("sgRNA", or simply "gNRA") can be engineered so as to incorporate aspects of both the crRNA and tracrRNA into a single RNA species – the guide RNA. See, *e.g.*, Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of which is hereby incorporated by reference. Cas9 recognizes a short motif in the CRISPR repeat sequences (the PAM or protospacer adjacent motif) to help distinguish self versus non-self. CRISPR biology, as well as Cas9 nuclease sequences and structures are well known to those of skill in the art (see, *e.g.*, "Complete genome sequence of an M1 strain of Streptococcus pyogenes." Ferretti *et al.*, J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011); and "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of each of which are incorporated herein by reference). Cas9 orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the organisms and loci disclosed in Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) RNA Biology 10:5, 726-737; the entire contents of which are incorporated herein by reference.

[177] In certain types of CRISPR systems (e.g., type II CRISPR systems), correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (rnc),

and a Cas9 protein. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves a linear or circular nucleic acid target complementary to the RNA. Specifically, the target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3′-5′ exonucleolytically. In nature, DNA-binding and cleavage typically requires protein and both RNAs. However, single guide RNAs ("sgRNA", or simply "gRNA") can be engineered to incorporate embodiments of both the crRNA and tracrRNA into a single RNA species—the guide RNA.

[178] In general, a "CRISPR system" refers collectively to transcripts and other elements involved in the expression of or directing the activity of CRISPR-associated ("Cas") genes, including sequences encoding a Cas gene, a tracr (trans-activating CRISPR) sequence (e.g. tracrRNA or an active partial tracrRNA), a tracr mate sequence (encompassing a "direct repeat" and a tracrRNA-processed partial direct repeat in the context of an endogenous CRISPR system), a guide sequence (also referred to as a "spacer" in the context of an endogenous CRISPR system), or other sequences and transcripts from a CRISPR locus. The tracrRNA of the system is complementary (fully or partially) to the tracr mate sequence present on the guide RNA.

DNA synthesis template

[179] As used herein, the term "DNA synthesis template" refers to the region or portion of the extension arm of a PEgRNA that is utilized as a template strand by a polymerase of a prime editor to encode a 3′ single-strand DNA flap that contains the desired edit and which then, through the mechanism of prime editing, replaces the corresponding endogenous strand of DNA at the target site. The extension arm, including the DNA synthesis template, may be comprised of DNA or RNA. In the case of RNA, the polymerase of the prime editor can be an RNA-dependent DNA polymerase (e.g., a reverse transcriptase). In the case of DNA, the polymerase of the prime editor can be a DNA-dependent DNA polymerase. In various embodiments the DNA synthesis template may comprise the "edit template" and the "homology arm", and all or a portion of the optional 5′ end modifier region, e2. That is, depending on the nature of the e2 region (e.g., whether it includes a hairpin, toeloop, or stem/loop secondary structure), the polymerase may encode none, some, or all of the e2 region as well. Said another way, in the case of a 3′ extension arm, the DNA synthesis template can include the portion of the extension arm that spans from the 5′ end of the primer binding site (PBS) to 3′ end of the gRNA core that may operate as a template for the synthesis of a single-strand of DNA by a polymerase (e.g., a reverse

transcriptase). In the case of a 5′ extension arm, the DNA synthesis template can include the portion of the extension arm that spans from the 5′ end of the PEgRNA molecule to the 3′ end of the edit template. In some embodiments, the DNA synthesis template excludes the primer binding site (PBS) of PEgRNAs either having a 3′ extension arm or a 5′ extension arm. Certain embodiments described here refer to an "an RT template," which is inclusive of the edit template and the homology arm, i.e., the sequence of the PEgRNA extension arm which is actually used as a template during DNA synthesis. The term "RT template" is equivalent to the term "DNA synthesis template." In certain embodiments, an RT template may be used to refer to a template polynucleotide for reverse transcription, e.g., in a prime editing system, complex, or method using a prime editor having a polymerase that is a reverse transcriptase. In some embodiments, a DNA synthesis template may be used to refer to a template polynucleotide for DNA polymerization, e.g., RNA-dependent DNA polymerization or DNA-dependent DNA polymerization, e.g., in a prime editing system, complex, or method using a prime editor having a polymerase that is an RNA-dependent DNA polymerase or a DNA-dependent DNA polymerase.

[180] In some embodiments, the DNA synthesis template is a single-stranded portion of the PEgRNA that is 5' of the PBS and comprises a region of complementarity to the PAM strand (i.e., the non-target strand or the edit strand), and comprises one or more nucleotide edits compared to the endogenous sequence of the double stranded target DNA. In some embodiments, the DNA synthesis template is complementary or substantially complementary to a sequence on the non-target strand that is downstream of a nick site, except for one or more non-complementary nucleotides at the intended nucleotide edit positions. In some embodiments, the DNA synthesis template is complementary or substantially complementary to a sequence on the non-target strand that is immediately downstream (i.e., directly downstream) of a nick site, except for one or more non-complementary nucleotides at the intended nucleotide edit positions. In some embodiments, one or more of the non-complementary nucleotides at the intended nucleotide edit positions are immediately downstream of a nick site. In some embodiments, the DNA synthesis template comprises one or more nucleotide edits relative to the double-stranded target DNA sequence. In some embodiments, the DNA synthesis template comprises one or more nucleotide edits relative to the non-target strand of the double-stranded target DNA sequence. For each PEgRNA described herein, a nick site is characteristic of the particular

napDNAbp to which the gRNA core of the PEgRNA associates, and is characteristic of the particular PAM required for recognition and function of the napDNAbp. For example, for a PEgRNA that comprises a gRNA core that associates with a SpCas9, the nick site in the phosphodiester bond between bases three ("-3" position relative to the position 1 of the PAM sequence) and four ("-4" position relative to position 1 of the PAM sequence). In some embodiments, the DNA synthesis template and the primer binding site are immediately adjacent to each other. The terms "nucleotide edit", "nucleotide change", "desired nucleotide change", and "desired nucleotide edit" are used interchangeably to refer to a specific nucleotide edit, e.g., a specific deletion of one or more nucleotides, a specific insertion of one or more nucleotides, a specific substitution (or multiple substitutions) of one or more nucleotides, or a combination thereof, at a specific position in a DNA synthesis template of a PEgRNA to be incorporated in a target DNA sequence. In some embodiments, the DNA synthesis template comprises more than one nucleotide edit relative to the double-stranded target DNA sequence. In such embodiments, each nucleotide edit is a specific nucleotide edit at a specific position in the DNA synthesis template, each nucleotide edit is at a different specific position relative to any of the other nucleotide edits in the DNA synthesis template, and each nucleotide edit is independently selected from a specific deletion of one or more nucleotides, a specific insertion of one or more nucleotides, a specific substitution (or multiple substitutions) of one or more nucleotides, or a combination thereof. A nucleotide edit may refer to the edit on the DNA synthesis template as compared to the sequence on the target strand of the target gene, or a nucleotide edit may refer to the edit encoded by the DNA synthesis template on the newly synthesized single stranded DNA that replaces the endogenous target DNA sequence on the non-target strand.

Dominant Negative Variant

[181] The terms "dominant negative variant" and "dominant negative mutant" refer to genes or gene products (*e.g.*, proteins) that comprise a mutation that results in the gene product acting antagonistically to the wild-type gene product (*i.e.*, inhibiting its activity). Dominant negative mutations generally result in an altered molecular function (often inactive). For example, the present disclosure provides dominant negative variants of MMR proteins that inhibit the activity of wild-type MMR proteins (*e.g.*, the dominant negative MLH1 proteins described herein).

Edit template

[182] The term "edit template" refers to a portion of the extension arm that encodes the desired edit in the single strand 3′ DNA flap that is synthesized by the polymerase, e.g., a DNA-dependent DNA polymerase, RNA-dependent DNA polymerase (e.g., a reverse transcriptase). Certain embodiments described here refer to "an RT template," which refers to both the edit template and the homology arm together, i.e., the sequence of the PEgRNA extension arm which is actually used as a template during DNA synthesis. The term "RT edit template" is also equivalent to the term "DNA synthesis template," but wherein the RT edit template reflects the use of a prime editor having a polymerase that is a reverse transcriptase, and wherein the DNA synthesis template reflects more broadly the use of a prime editor having any polymerase.

Extension arm

[183] The term "extension arm" refers to a nucleotide sequence component of a PEgRNA which comprises a primer binding site and a DNA synthesis template (e.g., an edit template and a homology arm) for a polymerase (e.g., a reverse transcriptase). In some embodiments, the extension arm is located at the 3′ end of the guide RNA. In other embodiments, the extension arm is located at the 5′ end of the guide RNA. In some embodiments, the extension arm comprises a DNA synthesis template and a primer binding site. In some embodiments, the extension arm comprises the following components in a 5′ to 3′ direction: the DNA synthesis template and the primer binding site. In some embodiments, the extension arm also includes a homology arm. In various embodiments, the extension arm comprises the following components in a 5′ to 3′ direction: the homology arm, the edit template, and the primer binding site. Since polymerization activity of the reverse transcriptase is in the 5′ to 3′ direction, the preferred arrangement of the homology arm, edit template, and primer binding site is in the 5′ to 3′ direction such that the reverse transcriptase, once primed by an annealed primer sequence, polymerizes a single strand of DNA using the edit template as a complementary template strand.

[184] The extension arm may also be described as comprising generally two regions: a primer binding site (PBS) and a DNA synthesis template, for instance. The primer binding site binds to the primer sequence that is formed from the endogenous DNA strand of the target site when it becomes nicked by the prime editor complex, thereby exposing a 3′ end on the endogenous nicked strand. As explained herein, the binding of the primer sequence to the primer binding site on the extension arm of the PEgRNA creates a duplex region with an exposed 3′ end (i.e., the 3′

of the primer sequence), which then provides a substrate for a polymerase to begin polymerizing a single strand of DNA from the exposed 3′ end along the length of the DNA synthesis template. The sequence of the single strand DNA product is the complement of the DNA synthesis template. Polymerization continues towards the 5′ of the DNA synthesis template (or extension arm) until polymerization terminates. Thus, the DNA synthesis template represents the portion of the extension arm that is encoded into a single strand DNA product (i.e., the 3′ single strand DNA flap containing the desired genetic edit information) by the polymerase of the prime editor complex and which ultimately replaces the corresponding endogenous DNA strand of the target site that sits immediately downstream of the PE-induced nick site. Without being bound by theory, polymerization of the DNA synthesis template continues towards the 5′ end of the extension arm until a termination event. Polymerization may terminate in a variety of ways, including, but not limited to (a) reaching a 5′ terminus of the PEgRNA (e.g., in the case of the 5′ extension arm wherein the DNA polymerase simply runs out of template), (b) reaching an impassable RNA secondary structure (e.g., hairpin or stem/loop), or (c) reaching a replication termination signal, e.g., a specific nucleotide sequence that blocks or inhibits the polymerase, or a nucleic acid topological signal, such as, supercoiled DNA or RNA.

Fusion protein

[185] The term "fusion protein" as used herein refers to a hybrid polypeptide which comprises protein domains from at least two different proteins. One protein may be located at the amino-terminal (N-terminal) portion of the fusion protein or at the carboxy-terminal (C-terminal) protein thus forming an "amino-terminal fusion protein" or a "carboxy-terminal fusion protein," respectively. A protein may comprise different domains, for example, a nucleic acid binding domain (*e.g.*, the gRNA binding domain of Cas9 that directs the binding of the protein to a target site) and a nucleic acid cleavage domain or a catalytic domain of a nucleic-acid editing protein. Another example includes a Cas9 or equivalent thereof to a reverse transcriptase. Any of the proteins provided herein may be produced by any method known in the art. For example, the proteins provided herein may be produced via recombinant protein expression and purification, which is especially suited for fusion proteins comprising a peptide linker. Methods for recombinant protein expression and purification are well known, and include those described by Green and Sambrook, *Molecular Cloning: A Laboratory Manual* (4th ed., Cold Spring Harbor

Laboratory Press, Cold Spring Harbor, N.Y. (2012)), the entire contents of which are incorporated herein by reference.

Guide RNA ("gRNA")

[186] As used herein, the term "guide RNA" is a particular type of guide nucleic acid which is mostly commonly associated with a Cas protein of a CRISPR-Cas9 and which associates with Cas9, directing the Cas9 protein to a specific sequence in a DNA molecule that includes complementarity to the protospacer sequence of the guide RNA. However, this term also embraces the equivalent guide nucleic acid molecules that associate with Cas9 equivalents, homologs, orthologs, or paralogs, whether naturally occurring or non-naturally occurring (e.g., engineered or recombinant), and which otherwise program the Cas9 equivalent to localize to a specific target nucleotide sequence. The Cas9 equivalents may include other napDNAbp from any type of CRISPR system (e.g., type II, V, VI), including Cpf1 (a type-V CRISPR-Cas systems), C2c1 (a type V CRISPR-Cas system), C2c2 (a type VI CRISPR-Cas system) and C2c3 (a type V CRISPR-Cas system). Further Cas-equivalents are described in Makarova *et al.*, "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector," Science 2016; 353(6299), the contents of which are incorporated herein by reference. Exemplary sequences are and structures of guide RNAs are provided herein. In addition, methods for designing appropriate guide RNA sequences are provided herein. As used herein, the "guide RNA" may also be referred to as a "traditional guide RNA" to contrast it with the modified forms of guide RNA termed "prime editing guide RNAs" (or "PEgRNAs").

[187] Guide RNAs or PEgRNAs may comprise various structural elements that include, but are not limited to:

[188] Spacer sequence – the sequence in the guide RNA or PEgRNA (having about 20 nts in length) which binds to the protospacer in the target DNA.

[189] gRNA core (or gRNA scaffold or backbone sequence) - refers to the sequence within the gRNA that is responsible for Cas9 binding, it does not include the 20 bp spacer/targeting sequence that is used to guide Cas9 to target DNA. In some embodiments, the gRNA core or scaffold comprises a sequence that comprises one or more nucleotide alterations compared to a naturally occurring CRISPR-Cas guide RNA scaffold, for example, a Cas9 guide RNA scaffold. In some embodiments, the sequence of the gRNA core is designed to comprise minimal or no sequence homology to the endogenous sequence of the target nucleic acid at the target site,

thereby reducing unintended edits. For example, in some embodiments, one or more base pairs in the second stem loop of a Cas9 gRNA core may be "flipped" (e.g., the G-U base pair and the U-A base pair as exemplified in Fig. 49A) to reduce unintended edits. In some embodiments, the gRNA core comprises no more than 1%, 5%, 10%, 15%, 20%, 25%, or 30% sequence homology to the sequence of the double stranded target DNA that flanks 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 nucleotides upstream or downstream of the position of the one or more nucleotide edits

[190] Extension arm – a single strand extension at the 3′ end or the 5′ end of the PEgRNA which comprises a primer binding site and a DNA synthesis template sequence that encodes via a polymerase (e.g., a reverse transcriptase) a single stranded DNA flap containing the genetic change of interest, which then integrates into the endogenous DNA by replacing the corresponding endogenous strand, thereby installing the desired genetic change.

[191] Transcription terminator – the guide RNA or PEgRNA may comprise a transcriptional termination sequence at the 3′ of the molecule. In some embodiments, the PEgRNA comprises a transcriptional termination sequence between the DNA synthesis template and the gRNA core.

Homology

[192] The terms "homologous," "homology," or "percent homology" as used herein refer to the degree of sequence identity between an amino acid or polynucleotide sequence and a corresponding reference sequence. "Homology" can refer to polymeric sequences, *e.g.*, polypeptide or DNA sequences that are similar. Homology can mean, for example, nucleic acid sequences with at least about: 50%, 55%, 60%, 65%, 70%, 75%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or 100% identity. In other embodiments, a "homologous sequence" of nucleic acid sequences may exhibit 93%, 95%, or 98% sequence identity to the reference nucleic acid sequence. For example, a "region of homology to a genomic region" can be a region of DNA that has a similar sequence to a given genomic region in the genome. A region of homology can be of any length that is sufficient to promote binding of a spacer or protospacer sequence to the genomic region. For example, the region of homology can comprise at least 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, or more bases in length such

that the region of homology has sufficient homology to undergo binding with the corresponding genomic region.

[193] When a percentage of sequence homology or identity is specified, in the context of two nucleic acid sequences or two polypeptide sequences, the percentage of homology or identity generally refers to the alignment of two or more sequences across a portion of their length when compared and aligned for maximum correspondence. When a position in the compared sequence can be occupied by the same base or amino acid, then the molecules can be homologous at that position. Unless stated otherwise, sequence homology or identity is assessed over the specified length of the nucleic acid, polypeptide, or portion thereof. In some embodiments, the homology or identity is assessed over a functional portion or a specified portion of the length.

[194] Alignment of sequences for assessment of sequence homology can be conducted by algorithms known in the art, such as the Basic Local Alignment Search Tool (BLAST) algorithm, which is described in Altschul *et al*, *J. Mol. Biol.* 215:403- 410, 1990. A publicly available, internet interface, for performing BLAST analyses is accessible through the National Center for Biotechnology Information. Additional known algorithms include those published in: Smith & Waterman, "Comparison of Biosequences", *Adv. Appl. Math.* 2:482, 1981; Needleman & Wunsch, "A general method applicable to the search for similarities in the amino acid sequence of two proteins" *J. Mol. Biol.* 48:443, 1970; Pearson & Lipman "Improved tools for biological sequence comparison", *Proc. Natl. Acad. Sci. USA* 85:2444, 1988; or by automated implementation of these or similar algorithms. Global alignment programs may also be used to align similar sequences of roughly equal size. Examples of global alignment programs include NEEDLE (available at www.ebi.ac uk/Tools/psa/emboss_needle/) which is part of the EMBOSS package (Rice P *et al.*, *Trends Genet.*, 2000; 16: 276-277), and the GGSEARCH program fasta.bioch.virginia.edu/fasta_www2/, which is part of the FASTA package (Pearson W and Lipman D, 1988, *Proc. Natl. Acad. Sci. USA*, 85: 2444-2448). Both of these programs are based on the Needleman-Wunsch algorithm, which is used to find the optimum alignment (including gaps) of two sequences along their entire length. A detailed discussion of sequence analysis can also be found in Unit 19.3 of Ausubel *et al* ("Current Protocols in Molecular Biology" John Wiley & Sons Inc, 1994-1998, Chapter 15, 1998).

[195] A skilled person understands that amino acid (or nucleotide) positions may be determined in homologous sequences based on alignment. For example, "H840" in a reference Cas9

sequence may correspond to H839, or another corresponding position in a Cas9 homolog when aligned to the reference Cas9 sequence.

Host cell

[196] The term "host cell," as used herein, refers to a cell that can host, replicate, and express a vector described herein, e.g., a vector comprising a nucleic acid molecule encoding an MLH1 variant and a fusion protein comprising a Cas9 or Cas9 equivalent and a reverse transcriptase.

Inhibit

[197] As used herein the term "inhibiting," "inhibit," or "inhibition" in the context of proteins and enzymes, for example, in the context of enzymes involved in the DNA mismatch repair pathway, refers to a reduction in the activity of the protein or enzyme. In some embodiments, the term refers to a reduction of the level of enzyme activity, e.g., the activity of one or more enzymes in the DNA mismatch repair pathway, to a level that is statistically significantly lower than an initial level, which may, for example, be a baseline level of enzyme activity. In some embodiments, the term refers to a reduction of the level of enzyme activity, e.g., the activity of one or more enzymes in the DNA mismatch repair pathway, to a level that is less than 75%, less than 50%, less than 40%, less than 30%, less than 25%, less than 20%, less than 10%, less than 9%, less than 8%, less than 7%, less than 6%, less than 5%, less than 4%, less than 3%, less than 2%, less than 1%, less than 0.5%, less than 0.1%, less than 0.01%, less than 0.001%, or less than 0.0001% of an initial level, which may, for example, be a baseline level of enzyme activity.

Linker

[198] The term "linker," as used herein, refers to a molecule linking two other molecules or moieties. The linker can be an amino acid sequence in the case of a linker joining two fusion proteins. For example, a Cas9 can be fused to a reverse transcriptase by an amino acid linker sequence. The linker can also be a nucleotide sequence in the case of joining two nucleotide sequences together. For example, in the instant case, the traditional guide RNA is linked via a spacer or linker nucleotide sequence to the RNA extension of a prime editing guide RNA which may comprise a RT template sequence and an RT primer binding site. In other embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 5-100 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80,

80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated. In certain embodiments, the linker is a self-hydrolyzing linker (*e.g.*, a 2A self-cleaving peptide as described further herein). Self-hydrolyzing linkers such as 2A self-cleaving peptides are capable of inducing ribosomal skipping during protein translation, resulting in the ribosome failing to make a peptide bond between two genes, or gene fragments.

MLH1

[199] The term "MLH1" refers to a gene encoding MLH1 (or MutL Homolog 1), a DNA mismatch repair enzyme. The protein encoded by this gene can heterodimerize with mismatch repair endonuclease PMS2 to form MutL alpha (MutLα), part of the DNA mismatch repair system. MLH1 mediates protein-protein interactions during mismatch recognition, strand discrimination, and strand removal. In mismatch repair, the heterodimer MSH2:MSH6 (MutSα) forms and binds the mismatch. MLH1 then forms a heterodimer with PMS2 (MutLα) and binds the MSH2:MSH6 heterodimer. The MutLα heterodimer then incises the nicked strand 5′ and 3′ of the mismatch, followed by excision of the mismatch from MutLα-generated nicks by EXO1. Finally, POLδ resynthesizes the excised strand, followed by LIG1 ligation.

[200] An exemplary amino acid sequence of MLH1 is human isoform 1, P40692-1: >sp|P40692|MLH1_HUMAN DNA mismatch repair protein Mlh1 OS=Homo sapiens OX=9606 GN=MLH1 PE=1 SV=1:

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL
DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN
QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR
SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF
LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMVKST
TSLTSSSTSGSSDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELP
APAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSV
LSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLS
EPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPL
EGLPIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKA
LRSHILPPKHFTEDGNILQLANLPDLYKVFERC (SEQ ID NO: 204),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

[201] Another exemplary amino acid sequence of MLH1 is human isoform 2, P40692-2 (wherein amino acids 1-241 of isoform 1 are missing): >sp|P40692-2|MLH1_HUMAN Isoform 2 of DNA mismatch repair protein Mlh1 OS=Homo sapiens OX=9606 GN=MLH1:

MNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHE
VHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMVKSTTSLTSSSTSGSSDKVYAHQMVR
TDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELPAPAEVAAKNQSLEGDTTKGTSE
MSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNH
SFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTE
EDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKEC
FESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFTEDGNILQLAN
LPDLYKVFERC (SEQ ID NO: 205),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 205.

[202] Another exemplary amino acid sequence of MLH1 is human isoform 3, P40692-3 (where amino acids 1-101 (MSFVAGVIRR…ASISTYGFRG (SEQ ID NO: 206) is replaced with MAF): >sp|P40692-2|MLH1_HUMAN Isoform 2 of DNA mismatch repair protein Mlh1 OS=Homo sapiens OX=9606 GN=MLH1:

MAFEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKALKNP
SEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFGNAVSRELIEIGCEDKTLAFKM
NGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVH
FLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMVKSTTSLTSSSTSGSSDKVYAHQMVRTD
SREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELPAPAEVAAKNQSLEGDTTKGTSEMS
EKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSF
VGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEED
GPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFES
LSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPD
LYKVFERC (SEQ ID NO: 207),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 207.

[203] The present disclosure contemplates targeting MLH1 and/or MMR pathway components that interact with MLH1, including any wildtype or naturally occurring variant of MLH1, including any amino acid sequence having at least 70%, or 75%, or 80%, or 85%, or 90%, or 95%, or 99% or more sequence identity with any of SEQ ID NOs: 204, 208-213, 215, 216, 218, 222, or 223, or nucleic acid molecules encoding any MLH1 or variant of MLH1 (*e.g.*, a dominant negative mutant of MLH1 as described herein), for inhibiting, blocking, or otherwise inactivating the wild type MLH1 function in the MMR pathway, and consequently, inhibiting, blocking, or otherwise inactivating the MMR pathway, e.g., during genome editing with a prime editor.

[204] In some embodiments, inactivation of the MMR pathway involves an inhibitor that disrupts, blocks, interferes with, or otherwise inactivates the wild type function of the MLH1 protein. In some embodiments, inactivation of the MMR pathway involves a mutant of the MLH1 protein, for example, contacting a target cell with a MLH1 mutant protein or expressing in a target cell an MLH1 mutant nucleic acid that encodes an MLH1 mutant protein. In some embodiments, the MLH1 mutant protein interferes with, and thereby inactivates, the function of a wild type MLH1 protein in the MMR pathway. In some embodiments, the MLH1 mutant is a dominant negative mutant. In some embodiments, the MLH mutant protein is capable of binding to an MLH1-interacting protein, for example, MutS.

[205] Without being bound by theory, MLH1 dominant negative mutants function by saturating binding of MutS, thereby blocking MutS-wild type MLH1 binding and interfering with the function of the wild type MLH1 protein in the MMR pathway.

[206] In various embodiments, the dominant negative MLH1 can include, for example, MLH1 E34A, which is based on SEQ ID NO: 222 and having the following amino acid sequence (underline and bolded to show the E34A mutation):

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIK**A**MIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL
DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN
QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR
SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF
LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMVKST
TSLTSSSTSGSSDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELP
APAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSV
LSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLS
EPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPL
EGLPIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKA
LRSHILPPKHFTEDGNILQLANLPDLYKVFERC (SEQ ID NO: 222),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 222.

[207] In various other embodiments, the dominant negative MLH1 can include, for example, MLH1 Δ756, which is based on SEQ ID NO: 208 and having the following amino acid sequence (underline and bolded to show the Δ756 mutation at the C terminus of the sequence):

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL
DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN
QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR
SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF
LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMVKST

TSLTSSSTSGSSDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELP
APAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSV
LSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLS
EPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPL
EGLPIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKA
LRSHILPPKHFTEDGNILQLANLPDLYKVFER[-](SEQ ID NO: 208),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 208 (wherein the [-] indicates deleted amino acid residue(s) relative to the parent or wildtype sequence).

[208] In still other embodiments, the dominant negative MLH1 can include, for example, MLH1 Δ754-Δ756, which is based on SEQ ID NO: 209 and having the following amino acid sequence (underline and bolded to show the Δ754-Δ756 mutation at the C terminus of the sequence):

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL
DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN
QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR
SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF
LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMVKST
TSLTSSSTSGSSDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELP
APAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSV
LSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLS
EPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPL
EGLPIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKA
LRSHILPPKHFTEDGNILQLANLPDLYKVF[- - -] (SEQ ID NO: 209),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 209 (wherein the [- - -] indicates deleted amino acid residue(s) relative to the parent or wildtype sequence).

[209] In yet other embodiments, the dominant negative MLH1 can include, for example, MLH1 E34A Δ754-Δ756, which is based on SEQ ID NO: 210 and having the following amino acid sequence (underline and bolded to show the E34A and Δ754-Δ756 mutations):

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIK**A**MIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL
DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN
QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR
SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF
LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLLGSNSSRMYFTQTLLPGLAGPSGEMVKST
TSLTSSSTSGSSDKVYAHQMVRTDSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELP
APAEVAAKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSV
LSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLS
EPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPL

EGLPIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKA
LRSHILPPKHFTEDGNILQLANLPDLYKVF[- - -] (SEQ ID NO: 210),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 210.

[210] In certain embodiments, the dominant negative MLH1 can include, for example, MLH1 1-335, which is based on SEQ ID NO: 211 and having the following amino acid sequence (contains amino acids 1-335 of SEQ NO: 204):

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL
DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN
QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR
SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF
LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLL (SEQ ID NO: 211),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 211.

[211] In other embodiments, the dominant negative MLH1 can include, for example, MLH1 1-335 E34A, which is based on SEQ ID NO: 212 and having the following amino acid sequence (contains amino acids 1-335 of SEQ NO: 204 and a E34A mutation relative to SEQ ID NO: 204):

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIK**A**MIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL
DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN
QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR
SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF
LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLL (SEQ ID NO: 212),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 212.

[212] In still other embodiments, the dominant negative MLH1 can include, for example, MLH1 1-335 NLS^{SV40} (or referred to as $MLH1dn^{NTD}$, which is based on SEQ ID NO: 204 and having the following amino acid sequence (contains amino acids 1-335 of SEQ NO: 204 and an NLS sequence of SV40):

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL
DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN
QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR

SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLL**PKKKRKV** (SEQ ID NO: 213),

with the underlined and bolded portion referring to the NLS of SV40), or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 213.

[213] In still other embodiments, the dominant negative MLH1 can include, for example, MLH1 1-335 $NLS^{alternate}$ (which is based on SEQ ID NO: 204 and having the following amino acid sequence (contains amino acids 1-335 of SEQ NO: 204 and an alternate NLS sequence)):

MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGGLKLIQIQDNGTGIRKEDL DIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAHVTITTKTADGKCAYRASYSDGKLKAPPKPCAGN QGTQITVEDLFYNIATRRKALKNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIR SIFGNAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIETVYAAYLPKNTHPF LYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIESKLL-[alternate NLS sequence] (SEQ ID NO: 214)-[alternate NLS sequence],

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 214. The alternate NLS sequence can be any suitable NLS sequence, including but not limited to:

Description	**Sequence**	**SEQ ID NO:**
NLS	MKRTADGSEFESPKKKRKV	SEQ ID NO: 101
NLS	MDSLLMNRRKFLYQFKNVRWAKGRRETYLC	SEQ ID NO: 1
NLS of nucleoplasmin	AVKRPAATKKAGQAKKKKLD	SEQ ID NO: 133
NLS of EGL-13	MSRRRKANPTKLSENAKKLAKEVEN	SEQ ID NO: 134
NLS of c-MYC	PAAKRVKLD	SEQ ID NO: 135
NLS of TUS-protein	KLKIKRPVK	SEQ ID NO: 136
NLS of polyoma large T-Ag	VSRKRPRP	SEQ ID NO: 137
NLS of Hepatitis D virus antigen	EGAPPAKRAR	SEQ ID NO: 138
NLS of murine p53	PPQPKKKPLDGE	SEQ ID NO: 139
NLS of PE1 and PE2	SGGSKRTADGSEFEPKKKRKV	SEQ ID NO: 103

In some embodiments, an NLS sequence is appended to the N-terminus of a protein and begins with a methionine ("M"). In other embodiments, an NLS sequence may be appended at the C-terminus of a protein, or between multiple domains of a fusion protein, and does not begin with a methionine (i.e., the M in, for example, SEQ ID NOs: 101, 1, and 134 is not included in the NLS when it is appended at the C-terminus or between two domains in a fusion protein).

[214] In still other embodiments, the dominant negative MLH1 can include, for example, MLH1 501-756, which corresponds to a C-terminal fragment of SEQ ID NO: 204 that corresponds to amino acids 501-756 of SEQ ID NO: 204:

INLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFG
VLRLSEPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDN
YVPPLEGLPIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHI
VYKALRSHILPPKHFTEDGNILQLANLPDLYKVFERC (SEQ ID NO: 215),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 215.

[215] In still other embodiments, the dominant negative MLH1 can include, for example, MLH1 501-753, which corresponds to a C-terminal fragment of SEQ ID NO: 204 that corresponds to amino acids 501-753 of SEQ ID NO: 204:

INLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFG
VLRLSEPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDN
YVPPLEGLPIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHI
VYKALRSHILPPKHFTEDGNILQLANLPDLYKVF[- - -] (SEQ ID NO: 216),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 216.

[216] In still other embodiments, the dominant negative MLH1 can include, for example, MLH1 461-756, which is a C-terminal fragment of SEQ ID NO: 204 that corresponds to amino acids 461-756 of SEQ ID NO: 204:

KRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFV
GCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDG
PKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFESL
SKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPDL
YKVFERC (SEQ ID NO: 217),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 217.

[217] In various embodiments, the dominant negative MLH1 can include, for example, MLH1 461-753, which is a C-terminal fragment of SEQ ID NO: 204 that corresponds to amino acids 461-753 of SEQ ID NO: 204:

KRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFV
GCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDG
PKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFESL
SKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPDL
YKVF[- - -] (SEQ ID NO: 218),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 218.

In various other embodiments, the dominant negative MLH1 can include, for example, MLH1 461-753, which is a C-terminal fragment of SEQ ID NO: 204 that corresponds to amino acids 461-753 of SEQ ID NO: 204, and which further comprises an N-terminal NLS, e.g., NLS^{SV40}:

[NLS]-
KRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFV
GCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDG
PKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFESL
SKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPDL
YKVF[- - -] (SEQ ID NO: 218),

or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 218. The NLS sequence can be any suitable NLS sequence, including but not limited to SEQ ID NOs: 1, 101, 103, 133-139.

napDNAbp

[218] As used herein, the term "nucleic acid programmable DNA binding protein" or "napDNAbp," of which Cas9 is an example, refer to proteins that use RNA:DNA hybridization to target and bind to specific sequences in a DNA molecule. Each napDNAbp is associated with at least one guide nucleic acid (e.g., guide RNA), which localizes the napDNAbp to a DNA sequence that comprises a DNA strand (i.e., a target strand) that is complementary to the guide nucleic acid, or a portion thereof (e.g., the protospacer of a guide RNA). In other words, the guide nucleic-acid "programs" the napDNAbp (e.g., Cas9 or equivalent) to localize and bind to a complementary sequence.

[219] Without being bound by theory, the binding mechanism of a napDNAbp – guide RNA complex, in general, includes the step of forming an R-loop whereby the napDNAbp induces the unwinding of a double-strand DNA target, thereby separating the strands in the region bound by the napDNAbp. The guide RNA spacer sequence then hybridizes to the "target strand." This displaces a "non-target strand" that is complementary to the target strand, which forms the single

strand region of the R-loop. In some embodiments, the napDNAbp includes one or more nuclease activities, which then cut the DNA, leaving various types of lesions. For example, the napDNAbp may comprises a nuclease activity that cuts the non-target strand at a first location, and/or cuts the target strand at a second location. Depending on the nuclease activity, the target DNA can be cut to form a "double-stranded break" whereby both strands are cut. In other embodiments, the target DNA can be cut at only a single site, i.e., the DNA is "nicked" on one strand. Exemplary napDNAbp with different nuclease activities include "Cas9 nickase" ("nCas9") and a deactivated Cas9 having no nuclease activities ("dead Cas9" or "dCas9"). Exemplary sequences for these and other napDNAbp are provided herein.

Nickase

[220] The term "nickase," as used herein, may refer to a Cas9 with one of the two nuclease domains inactivated. This enzyme is capable of cleaving only one strand of a target DNA. As used herein, a "nickase" may refer to a napDNAbp (e.g., a Cas protein) which is capable of cleaving only one of the two complementary strands of a double-stranded target DNA sequence, thereby generating a nick in that strand. In some embodiments, the nickase cleaves a non-target strand of a double stranded target DNA sequence. In some embodiments, the nickase comprises an amino acid sequence with one or more mutations in a catalytic domain of a canonical napDNAbp (e.g., a Cas protein), wherein the one or more mutations reduces or abolishes nuclease activity of the catalytic domain. In certain embodiments, the napDNAbp is a Cas9 nickase, a Cas12a nickase, or a Cas12b1 nickase. In some embodiments, the nickase is a Cas9 that comprises one or more mutations in a RuvC-like domain relative to a wild type Cas9 sequence or to an equivalent amino acid position in other Cas9 variants or Cas9 equivalents. In some embodiments, the nickase is a Cas9 that comprises one or more mutations in an HNH-like domain relative to a wild type Cas9 sequence or to an equivalent amino acid position in other Cas9 variants or Cas9 equivalents. In some embodiments, the nickase is a Cas9 that comprises an aspartate-to-alanine substitution (D10A) in the RuvC I catalytic domain of Cas9 relative to a canonical Cas9 sequence or to an equivalent amino acid position in other Cas9 variants or Cas9 equivalents. In some embodiments, the nickase is a Cas9 that comprises a H840A, N854A, and/or N863A mutation relative to a canonical Cas9 sequence, or to an equivalent amino acid position in other Cas9 variants or Cas9 equivalents. In some embodiments, the term "Cas9 nickase" refers to a Cas9 with one of the two nuclease domains inactivated. This enzyme is

capable of cleaving only one strand of a target DNA. In some embodiments, the nickase is a Cas protein that is not a Cas9 nickase.

[221] In some embodiments, the napDNAbp of the prime editing complex comprises an endonuclease having nucleic acid programmable DNA binding ability. In some embodiments, the napDNAbp comprises an active endonuclease capable of cleaving both strands of a double stranded target DNA. In some embodiments, the napDNAbp is a nuclease active endonuclease, e.g., a nuclease active Cas protein, that can cleave both strands of a double stranded target DNA by generating a nick on each strand. For example, a nuclease active Cas protein can generate a cleavage (a nick) on each strand of a double stranded target DNA. In some embodiments, the two nicks on both strands are staggered nicks, for example, generated by a napDNAbp comprising a Cas12a or Cas12b1. In some embodiments, the two nicks on both strands are at the same genomic position, for example, generated by a napDNAbp comprising a nuclease active Cas9. In some embodiments, the napDNAbp comprises an endonuclease that is a nickase. For example, in some embodiments, the napDNAbp comprises an endonuclease comprising one or more mutations that reduce nuclease activity of the endonuclease, rendering it a nickase. In some embodiments, the napDNAbp comprises an inactive endonuclease. For example, in some embodiments, the napDNAbp comprises an endonuclease comprising one or more mutations that abolish the nuclease activity. In various embodiments, the napDNAbp is a Cas9 protein or variant thereof. The napDNAbp can also be a nuclease active Cas9, a nuclease inactive Cas9 (dCas9), or a Cas9 nickase (nCas9). In a preferred embodiment, the napDNAbp is Cas9 nickase (nCas9) that nicks only a single strand. In certain embodiments, the napDNAbp is a Cas9 nickase, a Cas12a nickase, or a Cas12b1 nickase. In some embodiments, the napDNAbp can be selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute, and optionally has a nickase activity such that only one strand is cut. In some embodiments, the napDNAbp is selected from Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute , and optionally has a nickase activity such that one DNA strand is cut preferentially to the other DNA strand.

Nick site

[222] The terms "cleavage site," "nick site," and "cut site" as used interchangeably herein in the context of prime editing, refer to a specific position in between two nucleotides or two base pairs in the double-stranded target DNA sequence. In some embodiments, the position of a nick site is determined relative to the position of a specific PAM sequence. In some embodiments, the nick site is the particular position where a nick will occur when the double stranded target DNA is contacted with a napDNAbp, e.g., a nickase such as a Cas nickase, that recognizes a specific PAM sequence. For each PEgRNA described herein, a nick site is characteristic of the particular napDNAbp to which the gRNA core of the PEgRNA associates with, and is characteristic of the particular PAM required for recognition and function of the napDNAbp. For example, for a PEgRNA that comprises a gRNA core that associates with a SpCas9, the nick site in the phosphodiester bond between bases three ("-3" position relative to the position 1 of the PAM sequence) and four ("-4" position relative to position 1 of the PAM sequence).

[223] In some embodiments, a nick site is in a target strand of the double-stranded target DNA sequence. In some embodiments, a nick site is in a non-target strand of the double-stranded target DNA sequence. In some embodiments, the nick site is in a protospacer sequence. In some embodiments, the nick site is adjacent to a protospacer sequence. In some embodiments, a nick site is downstream of a region, e.g., on a non-target strand, that is complementary to a primer binding site of a PEgRNA. In some embodiments, a nick site is downstream of a region, e.g., on a non-target strand, that binds to a primer binding site of a PEgRNA. In some embodiments, a nick site is immediately downstream of a region, e.g., on a non-target strand, that is complementary to a primer binding site of a PEgRNA. In some embodiments, the nick site is upstream of a specific PAM sequence on the non-target strand of the double stranded target DNA, wherein the PAM sequence is specific for recognition by a napDNAbp that associates with the gRNA core of a PEgRNA. In some embodiments, the nick site is downstream of a specific PAM sequence on the non-target strand of the double stranded target DNA. wherein the PAM sequence is specific for recognition by a napDNAbp that associates with the gRNA core of a PEgRNA. In some embodiments, the nick site is 3 nucleotides upstream of the PAM sequence, and the PAM sequence is recognized by a *Streptococcus pyogenes* Cas9 nickase, a *P. lavamentivorans* Cas9 nickase, a *C. diphtheriae* Cas9 nickase, a *N. cinerea* Cas9, a *S. aureus* Cas9, or a *N. lari* Cas9 nickase. In some embodiments, the nick site is 3 nucleotides upstream of

the PAM sequence, and the PAM sequence is recognized by a Cas9 nickase, wherein the Cas9 nickase comprises a nuclease active HNH domain and a nuclease inactive RuvC domain. In some embodiments, the nick site is 2 base pairs upstream of the PAM sequence, and the PAM sequence is recognized by a *S. thermophilus* Cas9 nickase.

Nucleic acid molecule

[224] The term "nucleic acid," as used herein, refers to a polymer of nucleotides. The polymer may include natural nucleosides (i.e., adenosine, thymidine, guanosine, cytidine, uridine, deoxyadenosine, deoxythymidine, deoxyguanosine, and deoxycytidine), nucleoside analogs (e.g., 2-aminoadenosine, 2-thiothymidine, inosine, pyrrolo-pyrimidine, 3-methyl adenosine, 5-methylcytidine, C5 bromouridine, C5 fluorouridine, C5 iodouridine, C5 propynyl uridine, C5 propynyl cytidine, C5 methylcytidine, 7 deazaadenosine, 7 deazaguanosine, 8 oxoadenosine, 8 oxoguanosine, O(6) methylguanine, 4-acetylcytidine, 5-(carboxyhydroxymethyl)uridine, dihydrouridine, methylpseudouridine, 1-methyl adenosine, 1-methyl guanosine, N6-methyl adenosine, and 2-thiocytidine), chemically modified bases, biologically modified bases (e.g., methylated bases), intercalated bases, modified sugars (e.g., 2'-fluororibose, ribose, 2'-deoxyribose, 2′-O-methylcytidine, arabinose, and hexose), or modified phosphate groups (e.g., phosphorothioates and 5′ N phosphoramidite linkages).

P53

[225] As used herein, "p53" refers to tumor protein 53. Among other functions, p53 plays a role in DNA damage and repair, specifically in its role in regulation of the cell cycle, apoptosis, and genomic stability. P53 can activate DNA repair proteins when DNA has been damaged. P53 may also arrest cell growth by holding the cell cycle at the G1/S regulation point on DNA damage recognition, thereby providing DNA repair proteins more time to fix the DNA damage before allowing the cell to continue the cell cycle. Thus, in some embodiments of the methods described herein, p53 is inhibited (*e.g.*, by the p53 inhibitor protein "i53," or another p53 inhibitor) to increase the efficiency of prime editing.

PEgRNA

[226] As used herein, the terms "prime editing guide RNA" or "PEgRNA" or "extended guide RNA" refer to a specialized form of a guide RNA that has been modified to include one or more additional sequences for implementing the prime editing methods and compositions described

herein. As described herein, the prime editing guide RNA comprise one or more "extended regions" of nucleic acid sequence. The extended regions may comprise, but are not limited to, single-stranded RNA or DNA. Further, the extended regions may occur at the 3′ end of a traditional guide RNA. In other arrangements, the extended regions may occur at the 5′ end of a traditional guide RNA. In still other arrangements, the extended region may occur at an intramolecular region of the traditional guide RNA, for example, in the gRNA core region which associates and/or binds to the napDNAbp. The extended region comprises a "DNA synthesis template" which encodes (by the polymerase of the prime editor) a single-stranded DNA which, in turn, has been designed to be (a) homologous with the endogenous target DNA to be edited, and (b) which comprises at least one desired nucleotide change (e.g., a transition, a transversion, a deletion, or an insertion) to be introduced or integrated into the endogenous target DNA. The extended region may also comprise other functional sequence elements, such as, but not limited to, a "primer binding site" and a "spacer or linker" sequence, or other structural elements, such as, but not limited to aptamers, stem loops, hairpins, toe loops (e.g., a 3′ toeloop), or an RNA-protein recruitment domain (e.g., MS2 hairpin). As used herein the "primer binding site" comprises a sequence that hybridizes to a single-strand DNA sequence having a 3′end generated from the nicked DNA of the R-loop.

[227] In certain embodiments, the PEgRNAs have a 5′ extension arm, a spacer, and a gRNA core. The 5′ extension further comprises in the 5′ to 3′ direction a reverse transcriptase template, a primer binding site, and a linker. The reverse transcriptase template may also be referred to more broadly as the "DNA synthesis template" where the polymerase of a prime editor described herein is not an RT, but another type of polymerase.

[228] In certain other embodiments, the PEgRNAs have a 5′ extension arm, a spacer, and a gRNA core. The 5′ extension further comprises in the 5′ to 3′ direction a reverse transcriptase template, a primer binding site, and a linker. The reverse transcriptase template may also be referred to more broadly as the "DNA synthesis template" where the polymerase of a prime editor described herein is not an RT, but another type of polymerase.

[229] In still other embodiments, the PEgRNAs have in the 5′ to 3′ direction a spacer (1), a gRNA core (2), and an extension arm (3). The extension arm (3) is at the 3′ end of the PEgRNA. The extension arm (3) further comprises in the 5′ to 3′ direction a "homology arm," an "edit template," and a "primer binding site." In certain embodiments, a PEgRNA comprises from 5' to

3', a space, a DNA synthesis template, and a primer binding site. The extension arm (3) may also comprise an optional modifier region at the 3′ and 5′ ends, which may be the same sequences or different sequences. In addition, the 3′ end of the PEgRNA may comprise a transcriptional terminator sequence. These sequence elements of the PEgRNAs are further described and defined herein.

[230] In still other embodiments, the PEgRNAs have in the 5′ to 3′ direction an extension arm (3), a spacer (1), and a gRNA core (2). The extension arm (3) is at the 5′ end of the PEgRNA. The extension arm (3) further comprises in the 3′ to 5′ direction a "homology arm," an "edit template," and a "primer binding site." The extension arm (3) may also comprise an optional modifier region at the 3′ and 5′ ends, which may be the same sequences or different sequences. The PEgRNAs may also comprise a transcriptional terminator sequence at the 3′ end. These sequence elements of the PEgRNAs are further described and defined herein.

PE1

[231] As used herein, "PE1" refers to a prime editor system comprising a fusion protein comprising Cas9(H840A) and a wild type MMLV RT having the following structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(wt)] + a desired PEgRNA, wherein the PE fusion has the amino acid sequence of SEQ ID NO: 100, which is shown as follows;

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA**
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPI
FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFI
QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN
FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKR
YDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV
KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY
HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL
KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKN
RGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK
HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA
LIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET
NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLP
KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY
LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR
YTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSS*TLNIE*
DEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLL
DQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFC
LRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATS

*ELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGT AGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGV LTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSN ARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSL LQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLL TSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP*SGGS KRTADGSEFEPKKKRKV (SEQ ID NO: 100)

KEY:

NUCLEAR LOCALIZATION SEQUENCE (NLS) TOP:(SEQ ID NO: 101), BOTTOM: (SEQ ID NO: 103)

CAS9(H840A) (SEQ ID NO: 37) (SEQ ID NO: 37 corresponds identically to SEQ ID NO: 2, except with an H840A substitution)

33-AMINO ACID LINKER (SEQ ID NO: 102)

M-MLV reverse transcriptase (SEQ ID NO: 81).

[232] Alternatively, PE1 may also refer to the prime editor fusion protein of SEQ ID NO: 100, i.e., without the pegRNA complexed thereto. PE1 may be complexed with a pegRNA during operation and/or use in prime editing.

PE2

[233] As used herein, "PE2" refers to a prime editing system comprising a fusion protein comprising Cas9(H840A) and a variant MMLV RT having the following structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)] + a desired PEgRNA, wherein the PE fusion has the amino acid sequence of SEQ ID NO: 107, which is shown as follows:

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPI FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFI QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKR YDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNS RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKN RGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLP KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR

YTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSS*TLNIE DEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLL DQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFC LRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSE LDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKA GFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLT QKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQ EGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTS EGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSGGSKR TADGSEFEPKKKRKV** (SEQ ID NO: 107)

KEY:

NUCLEAR LOCALIZATION SEQUENCE (NLS) TOP:(SEQ ID NO: 101), BOTTOM: (SEQ ID NO: 103)

CAS9(H840A) (SEQ ID NO: 37)

33-AMINO ACID LINKER (SEQ ID NO: 102)

M-MLV reverse transcriptase (SEQ ID NO: 98).

[234] Alternatively, PE2 may also refer to the prime editor fusion protein of SEQ ID NO: 107, i.e., without the pegRNA complexed thereto. PE2 may be complexed with a pegRNA during operation and/or use in prime editing.

PE3

[235] As used herein, "PE3" refers to PE2 plus a second-strand nicking guide RNA that complexes with the PE2 and introduces a nick in the non-edited DNA strand in order to induce preferential replacement of the edited strand.

PE3b

[236] As used herein, "PE3b" refers to PE3 but wherein the second-strand nicking guide RNA is designed for temporal control such that the second strand nick is not introduced until after the installation of the desired edit. This is achieved by designing a gRNA with a spacer sequence that matches only the edited strand, but not the original allele. Using this strategy, referred to hereafter as PE3b, mismatches between the protospacer and the unedited allele should disfavor nicking by the sgRNA until after the editing event on the PAM strand takes place.

PE4

[237] As used herein, "PE4" refers to a system comprising PE2 plus an MLH1 dominant negative protein (*i.e.*, wild-type MLH1 with amino acids 754-756 truncated as described further herein) expressed *in trans*. In some embodiments, PE4 refers to a fusion protein comprising PE2 and an MLH1 dominant negative protein joined via an optional linker.

PE5

[238] As used herein, "PE5" refers to a system comprising PE3 plus an MLH1 dominant negative protein (*i.e.*, wild-type MLH1 with amino acids 754-756 deleted as described further herein, which may be referred to as "MLH1 Δ754-756" or "MLH1dn") expressed *in trans*. In some embodiments, PE5 refers to a fusion protein comprising PE3 and an MLH1 dominant negative protein joined via an optional linker.

PEmax

[239] As used herein, "PEmax" (see FIG. 54B) refers to a PE complex comprising a fusion protein comprising Cas9(R221K N394K H840A) and a variant MMLV RT pentamutant (D200N T306K W313F T330P L603W) having the following structure: [bipartite NLS]-[Cas9(R221K)(N394K)(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)]-[bipartite NLS]-[NLS] + a desired PEgRNA, wherein the PE fusion has the amino acid sequence of SEQ ID NO: 99.

PE4max

[240] As used herein, "PE4max" refers to PE4 but wherein the PE2 component is substituted with PEmax.

PE5max

[241] As used herein, "PE5max" refers to PE5 but wherein the PE2 component of PE3 is substituted with PEmax.

PE-short

[242] As used herein, "PE-short" refers to a PE construct that is fused to a C-terminally truncated reverse transcriptase, and has the following amino acid sequence:

MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGA**
LLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPI
FGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFI
QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSN
FDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKR
YDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV
KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNS
RFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKV
KYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY
HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGR
LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL
KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKN
RGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK

HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTA LIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIET NGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLP KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHY LDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR YTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSS*TLNIE DEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLL DQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFC LRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSE LDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKA GFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLT QKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNAR MTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDNSRLIN**SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 117)

KEY:

NUCLEAR LOCALIZATION SEQUENCE (NLS) TOP:(SEQ ID NO: 101), BOTTOM: (SEQ ID NO: 103)

CAS9(H840A) (SEQ ID NO: 37)

33-AMINO ACID LINKER 1 (SEQ ID NO: 102)

M-MLV TRUNCATED reverse transcriptase (SEQ ID NO: 80)

Polymerase

[243] As used herein, the term "polymerase" refers to an enzyme that synthesizes a nucleotide strand and that may be used in connection with the prime editor systems described herein. The polymerase can be a "template-dependent" polymerase (i.e., a polymerase that synthesizes a nucleotide strand based on the order of nucleotide bases of a template strand). The polymerase can also be a "template-independent" polymerase (i.e., a polymerase that synthesizes a nucleotide strand without the requirement of a template strand). A polymerase may also be further categorized as a "DNA polymerase" or an "RNA polymerase." In various embodiments, the prime editor system comprises a DNA polymerase. In various embodiments, the DNA polymerase can be a "DNA-dependent DNA polymerase" (i.e., whereby the template molecule is a strand of DNA). In such cases, the DNA template molecule can be a PEgRNA, wherein the extension arm comprises a strand of DNA. In such cases, the PEgRNA may be referred to as a chimeric or hybrid PEgRNA which comprises an RNA portion (i.e., the guide RNA components, including the spacer and the gRNA core) and a DNA portion (i.e., the extension arm). In various other embodiments, the DNA polymerase can be an "RNA-dependent DNA polymerase" (i.e., whereby the template molecule is a strand of RNA). In such cases, the PEgRNA is RNA, i.e., including an RNA extension. The term "polymerase" may also refer to an enzyme that catalyzes the polymerization of nucleotide (i.e., the polymerase activity). Generally, the enzyme will

initiate synthesis at the 3'-end of a primer annealed to a polynucleotide template sequence (e.g., such as a primer sequence annealed to the primer binding site of a PEgRNA) and will proceed toward the 5′ end of the template strand. A "DNA polymerase" catalyzes the polymerization of deoxynucleotides. As used herein in reference to a DNA polymerase, the term DNA polymerase includes a "functional fragment thereof." A "functional fragment thereof" refers to any portion of a wild-type or mutant DNA polymerase that encompasses less than the entire amino acid sequence of the polymerase and which retains the ability, under at least one set of conditions, to catalyze the polymerization of a polynucleotide. Such a functional fragment may exist as a separate entity, or it may be a constituent of a larger polypeptide, such as a fusion protein.

Prime editing

[244] As used herein, the term "prime editing" refers to an approach for gene editing using napDNAbps, a polymerase (e.g., a reverse transcriptase), and specialized guide RNAs that include a DNA synthesis template for encoding desired new genetic information (or deleting genetic information) that is then incorporated into a target DNA sequence. Certain embodiments of prime editing are described in the embodiments of FIG. 1. Classical prime editing is described in the inventors' publication of Anzalone, A. V. *et al.* Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* **576**, 149–157 (2019), which is incorporated herein by reference in its entirety.

[245] Prime editing represents a platform for genome editing that is a versatile and precise genome editing method that directly writes new genetic information into a specified DNA site using a nucleic acid programmable DNA binding protein ("napDNAbp") working in association with a polymerase (i.e., in the form of a fusion protein or otherwise provided *in trans* with the napDNAbp), wherein the prime editing system is programmed with a prime editing (PE) guide RNA ("PEgRNA") that both specifies the target site and templates the synthesis of the desired edit in the form of a replacement DNA strand by way of an extension (either DNA or RNA) engineered onto a guide RNA (e.g., at the 5′ or 3′ end, or at an internal portion of a guide RNA). The replacement strand containing the desired edit (e.g., a single nucleobase substitution) shares the same (or is homologous to) sequence as the endogenous strand (immediately downstream of the nick site) of the target site to be edited (with the exception that it includes the desired edit). Through DNA repair and/or replication machinery, the endogenous strand downstream of the nick site is replaced by the newly synthesized replacement strand containing the desired edit. In

some cases, prime editing may be thought of as a "search-and-replace" genome editing technology since the prime editors, as described herein, not only search and locate the desired target site to be edited, but at the same time, encode a replacement strand containing a desired edit which is installed in place of the corresponding target site endogenous DNA strand. The prime editors of the present disclosure relate, in part, to the mechanism of target-primed reverse transcription (TPRT), which can be engineered for conducting precision CRISPR/Cas-based genome editing with high efficiency and genetic flexibility. TPRT is used by mobile DNA elements, such as mammalian non-LTR retrotransposons and bacterial Group II introns. The inventors have herein used Cas protein-reverse transcriptase fusions or related systems to target a specific DNA sequence with a guide RNA, generate a single strand nick at the target site, and use the nicked DNA as a primer for reverse transcription of an engineered reverse transcriptase template that is integrated with the guide RNA. However, while the concept begins with prime editors that use reverse transcriptase as the DNA polymerase component, the prime editors described herein are not limited to reverse transcriptases but may include the use of virtually any DNA polymerase. Indeed, while the application throughout may refer to prime editors with "reverse transcriptases," it is set forth here that reverse transcriptases are only one type of DNA polymerase that may work with prime editing. Thus, wherever the specification mentions a "reverse transcriptase," the person having ordinary skill in the art should appreciate that any suitable DNA polymerase may be used in place of the reverse transcriptase. Thus, in one aspect, the prime editors may comprise Cas9 (or an equivalent napDNAbp) which is programmed to target a DNA sequence by associating it with a specialized guide RNA (i.e., PEgRNA) containing a spacer sequence that anneals to a complementary protospacer in the target DNA. The specialized guide RNA also contains new genetic information in the form of an extension that encodes a replacement strand of DNA containing a desired genetic alteration which is used to replace a corresponding endogenous DNA strand at the target site. To transfer information from the PEgRNA to the target DNA, the mechanism of prime editing involves nicking the target site in one strand of the DNA to expose a 3′-hydroxyl group. The exposed 3′-hydroxyl group can then be used to prime the DNA polymerization of the edit-encoding extension on PEgRNA directly into the target site. In various embodiments, the extension—which provides the template for polymerization of the replacement strand containing the edit—can be formed from RNA or DNA. In the case of an RNA extension, the polymerase of the prime editor can be an RNA-

dependent DNA polymerase (such as, a reverse transcriptase). In the case of a DNA extension, the polymerase of the prime editor may be a DNA-dependent DNA polymerase. The newly synthesized strand (i.e., the replacement DNA strand containing the desired edit) that is formed by the herein disclosed prime editors would be homologous to the genomic target sequence (i.e., have the same sequence as) except for the inclusion of a desired nucleotide change (e.g., a single nucleotide change, a deletion, or an insertion, or a combination thereof). The newly synthesized (or replacement) strand of DNA may also be referred to as a single strand DNA flap, which would compete for hybridization with the complementary homologous endogenous DNA strand, thereby displacing the corresponding endogenous strand. In certain embodiments, the system can be combined with the use of an error-prone reverse transcriptase enzyme (e.g., provided as a fusion protein with the Cas9 domain, or provided *in trans* to the Cas9 domain). The error-prone reverse transcriptase enzyme can introduce alterations during synthesis of the single strand DNA flap. Thus, in certain embodiments, error-prone reverse transcriptase can be utilized to introduce nucleotide changes to the target DNA. Depending on the error-prone reverse transcriptase that is used with the system, the changes can be random or non-random. Resolution of the hybridized intermediate (comprising the single strand DNA flap synthesized by the reverse transcriptase hybridized to the endogenous DNA strand) can include removal of the resulting displaced flap of endogenous DNA (e.g., with a 5′ end DNA flap endonuclease, FEN1), ligation of the synthesized single strand DNA flap to the target DNA, and assimilation of the desired nucleotide change as a result of cellular DNA repair and/or replication processes. Because templated DNA synthesis offers single nucleotide precision for the modification of any nucleotide, including insertions and deletions, the scope of this approach is very broad and could foreseeably be used for myriad applications in basic science and therapeutics.

[246] In various embodiments, prime editing operates by contacting a target DNA molecule (for which a change in the nucleotide sequence is desired to be introduced) with a nucleic acid programmable DNA binding protein (napDNAbp) complexed with a prime editing guide RNA (PEgRNA). In various embodiments, the prime editing guide RNA (PEgRNA) comprises an extension at the 3′ or 5′ end of the guide RNA, or at an intramolecular location in the guide RNA and encodes the desired nucleotide change (e.g., single nucleotide change, insertion, or deletion). In step (a), the napDNAbp/extended gRNA complex contacts the DNA molecule, and the extended gRNA guides the napDNAbp to bind to a target locus. In step (b), a nick in one of the

strands of DNA of the target locus is introduced (e.g., by a nuclease or chemical agent), thereby creating an available 3′ end in one of the strands of the target locus. In certain embodiments, the nick is created in the strand of DNA that corresponds to the R-loop strand, i.e., the strand that is not hybridized to the guide RNA sequence, i.e., the "non-target strand." The nick, however, could be introduced in either of the strands. That is, the nick could be introduced into the R-loop "target strand" (i.e., the strand hybridized to the protospacer of the extended gRNA) or the "non-target strand" (i.e., the strand forming the single-stranded portion of the R-loop and which is complementary to the target strand). In step (c), the 3′ end of the DNA strand (formed by the nick) interacts with the extended portion of the guide RNA in order to prime reverse transcription (i.e., "target-primed RT"). In certain embodiments, the 3′ end DNA strand hybridizes to a specific RT priming sequence on the extended portion of the guide RNA, i.e., the "reverse transcriptase priming sequence" or "primer binding site" on the PEgRNA. In step (d), a reverse transcriptase (or other suitable DNA polymerase) is introduced which synthesizes a single strand of DNA from the 3′ end of the primed site towards the 5′ end of the prime editing guide RNA. The DNA polymerase (e.g., reverse transcriptase) can be fused to the napDNAbp or alternatively can be provided *in trans* to the napDNAbp. This forms a single-strand DNA flap comprising the desired nucleotide change (e.g., the single base change, insertion, or deletion, or a combination thereof) and which is otherwise homologous to the endogenous DNA at or adjacent to the nick site. In step (e), the napDNAbp and guide RNA are released. Steps (f) and (g) relate to the resolution of the single strand DNA flap such that the desired nucleotide change becomes incorporated into the target locus. This process can be driven towards the desired product formation by removing the corresponding 5′ endogenous DNA flap that forms once the 3′ single strand DNA flap invades and hybridizes to the endogenous DNA sequence. Without being bound by theory, the cells endogenous DNA repair and replication processes resolves the mismatched DNA to incorporate the nucleotide change(s) to form the desired altered product. The process can also be driven towards product formation with "second strand nicking." This process may introduce at least one or more of the following genetic changes: transversions, transitions, deletions, and insertions.

[247] The term "prime editor (PE) system" or "prime editor (PE)" or "PE system" or "PE editing system" refers the compositions involved in the method of genome editing using target-primed reverse transcription (TPRT) describe herein, including, but not limited to the

napDNAbps, reverse transcriptases (or another DNA polymerase), fusion proteins (e.g., comprising napDNAbps and reverse transcriptases or comprising napDNAbps and DNA polymerases), prime editing guide RNAs, and complexes comprising fusion proteins and prime editing guide RNAs, as well as accessory elements, such as second strand nicking components (e.g., second strand sgRNAs) and 5′ endogenous DNA flap removal endonucleases (e.g., FEN1) for helping to drive the prime editing process towards the edited product formation.

[248] Although in the embodiments described thus far the PEgRNA constitutes a single molecule comprising a guide RNA (which itself comprises a spacer sequence and a gRNA core or scaffold) and a 5′ or 3′ extension arm comprising the primer binding site and a DNA synthesis template, the PEgRNA may also take the form of two individual molecules comprised of a guide RNA and a *trans* prime editor RNA template (tPERT), which essentially houses the extension arm (including, in particular, the primer binding site and the DNA synthesis domain) and an RNA-protein recruitment domain (e.g., MS2 aptamer or hairpin) in the same molecule which becomes co-localized or recruited to a modified prime editor complex that comprises a tPERT recruiting protein (e.g., MS2cp protein, which binds to the MS2 aptamer).

Prime editor

[249] The term "prime editor" refers to fusion constructs comprising a napDNAbp (e.g., Cas9 nickase) and a reverse transcriptase and is capable of carrying out prime editing on a target nucleotide sequence in the presence of a PEgRNA (or "extended guide RNA"). The term "prime editor" may refer to the fusion protein or to the fusion protein complexed with a PEgRNA, and/or further complexed with a second-strand nicking sgRNA. In some embodiments, the prime editor may also refer to the complex comprising a fusion protein (reverse transcriptase fused to a napDNAbp), a PEgRNA, and a regular guide RNA capable of directing the second-site nicking step of the non-edited strand as described herein. In certain embodiments, a prime editor (e.g., PE1, PE2, or PE3) may be provided as a system along with an inhibitor of the DNA mismatch repair pathway, such as a dominant negative MLH1 protein. In various embodiments, the inhibitor of the DNA mismatch repair pathway, such as a dominant negative MLH1 protein, may be provided *in trans* to the prime editor. In other embodiments, the inhibitor of the DNA mismatch repair pathway, such as a dominant negative MLH1 protein, may be complexed to the prime editor, e.g., coupled through a linker to the prime editor fusion protein.

Primer binding site

[250] The term "primer binding site" or "the PBS" refers to the portion of a PEgRNA as a component of the extension arm (for example, at the 3′ end of the extension arm) . The term "primer binding site" refers to a single-stranded portion of the PEgRNA as a component of the extension arm that comprises a region of complementarity to a sequence on the non-target strand. In some embodiments, the primer binding site is complementary to a region upstream of a nick site in a non-target strand. In some embodiments, the primer binding site is complementary to a region immediately upstream of a nick site in the non-target strand. In some embodiments, the primer binding site is capable of binding to the primer sequence that is formed after nicking (e.g., by a nickase component of a prime editor, for example, a Cas9 nikcase) of the target sequence by the prime editor. When a prime editor nicks one strand of the target DNA sequence (e.g., by a Cas nickase component of the prime editor), a 3′-ended ssDNA flap is formed, which serves a primer sequence that anneals to the primer binding site on the PEgRNA to prime reverse transcription. In some embodiments, the PBS is complementary to, or substantially complementary to, and can anneal to, a free 3' end on the non-target strand of the double stranded target DNA at the nick site. In some embodiments, the PBS annealed to the free 3' end on the non-target strand can initiate target-primed DNA synthesis.

Protospacer

[251] As used herein, the term "protospacer" refers to the sequence (~20 bp) in DNA adjacent to the PAM (protospacer adjacent motif) sequence. The protospacer shares the same sequence as the spacer sequence of the guide RNA. The guide RNA anneals to the complement of the protospacer sequence on the target DNA (specifically, one strand thereof, i.e., the "target strand" versus the "non-target strand" of the target DNA sequence). In some embodiments, in order for a Cas nickase component of the prime editor to function it also requires a specific protospacer adjacent motif (PAM) which varies depending on the Cas protein component itself, e.g., the type of Cas protein. For example, the most commonly used Cas9 nuclease, derived from *S. pyogenes*, recognizes a PAM sequence of NGG that is found directly downstream of the target sequence in the genomic DNA, on the non-target strand. The skilled person will appreciate that the literature in the state of the art sometimes refers to the "protospacer" as the ~20-nt target-specific guide sequence on the guide RNA itself, rather than referring to it as a "spacer." Thus, in some cases, the term "protospacer" as used herein may be used interchangeably with the term "spacer." The

context of the description surrounding the appearance of either "protospacer" or "spacer" will help inform the reader as to whether the term is in reference to the gRNA or the DNA target.

Protospacer adjacent motif (PAM)

[252] As used herein, the term "protospacer adjacent sequence" or "PAM" refers to an approximately 2-6 base pair DNA sequence that is an important targeting component of a Cas9 nuclease. Typically, the PAM sequence is on either strand, and is downstream in the 5′ to 3′ direction of the Cas9 cut site. The canonical PAM sequence (i.e., the PAM sequence that is associated with the Cas9 nuclease of *Streptococcus pyogenes* or SpCas9) is 5′-NGG-3′ wherein "N" is any nucleobase followed by two guanine ("G") nucleobases. Different PAM sequences can be associated with different Cas9 nucleases or equivalent proteins from different organisms. In addition, any given Cas9 nuclease, e.g., SpCas9, may be modified to alter the PAM specificity of the nuclease such that the nuclease recognizes alternative PAM sequence.

[253] For example, with reference to the canonical SpCas9 amino acid sequence is SEQ ID NO: 2, the PAM sequence can be modified by introducing one or more mutations, including (a) D1135V, R1335Q, and T1337R "the VQR variant", which alters the PAM specificity to NGAN or NGNG, (b) D1135E, R1335Q, and T1337R "the EQR variant", which alters the PAM specificity to NGAG, and (c) D1135V, G1218R, R1335E, and T1337R "the VRER variant", which alters the PAM specificity to NGCG. In addition, the D1135E variant of canonical SpCas9 still recognizes NGG, but it is more selective compared to the wild type SpCas9 protein.

[254] It will also be appreciated that Cas9 enzymes from different bacterial species (i.e., Cas9 orthologs) can have varying PAM specificities. For example, Cas9 from *Staphylococcus aureus* (SaCas9) recognizes NGRRT or NGRRN. In addition, Cas9 from *Neisseria meningitis* (NmCas) recognizes NNNNGATT. In another example, Cas9 from *Streptococcus thermophilis* (StCas9) recognizes NNAGAAW. In still another example, Cas9 from *Treponema denticola* (TdCas) recognizes NAAAAC. These are examples and are not meant to be limiting. It will be further appreciated that non-SpCas9s bind a variety of PAM sequences, which makes them useful when no suitable SpCas9 PAM sequence is present at the desired target cut site. Furthermore, non-SpCas9s may have other characteristics that make them more useful than SpCas9. For example, Cas9 from *Staphylococcus aureus* (SaCas9) is about 1 kilobase smaller than SpCas9, so it can be packaged into adeno-associated virus (AAV). Further reference may be made to Shah *et al.*,

"Protospacer recognition motifs: mixed identities and functional diversity," *RNA Biology*, 10(5): 891-899 (which is incorporated herein by reference).

Reverse transcriptase

[255] The term "reverse transcriptase" describes a class of polymerases characterized as RNA-dependent DNA polymerases. All known reverse transcriptases require a primer to synthesize a DNA transcript from an RNA template. Historically, reverse transcriptase has been used primarily to transcribe mRNA into cDNA which can then be cloned into a vector for further manipulation. Avian myoblastosis virus (AMV) reverse transcriptase was the first widely used RNA-dependent DNA polymerase (Verma, Biochim. Biophys. Acta 473:1 (1977)). The enzyme has 5′-3′ RNA-directed DNA polymerase activity, 5′-3′ DNA-directed DNA polymerase activity, and RNase H activity. RNase H is a processive 5′ and 3′ ribonuclease specific for the RNA strand for RNA-DNA hybrids (Perbal, A Practical Guide to Molecular Cloning, New York: Wiley & Sons (1984)). Errors in transcription cannot be corrected by reverse transcriptase because known viral reverse transcriptases lack the 3′-5′ exonuclease activity necessary for proofreading (Saunders and Saunders, Microbial Genetics Applied to Biotechnology, London: Croom Helm (1987)). A detailed study of the activity of AMV reverse transcriptase and its associated RNase H activity has been presented by Berger *et al.*, Biochemistry 22:2365-2372 (1983). Another reverse transcriptase which is used extensively in molecular biology is reverse transcriptase originating from Moloney murine leukemia virus (M-MLV). See, e.g., Gerard, G. R., DNA 5:271-279 (1986) and Kotewicz, M. L., *et al.*, Gene 35:249-258 (1985). M-MLV reverse transcriptase substantially lacking in RNase H activity has also been described. See, e.g., U.S. Pat. No. 5,244,797. The invention contemplates the use of any such reverse transcriptases, or variants or mutants thereof.

[256] In addition, the invention contemplates the use of reverse transcriptases that are error-prone, i.e., that may be referred to as error-prone reverse transcriptases or reverse transcriptases that do not support high fidelity incorporation of nucleotides during polymerization. During synthesis of the single-strand DNA flap based on the RT template integrated with the guide RNA, the error-prone reverse transcriptase can introduce one or more nucleotides which are mismatched with the RT template sequence, thereby introducing changes to the nucleotide sequence through erroneous polymerization of the single-strand DNA flap. These errors introduced during synthesis of the single strand DNA flap then become integrated into the double

strand molecule through hybridization to the corresponding endogenous target strand, removal of the endogenous displaced strand, ligation, and then through one more round of endogenous DNA repair and/or sequencing processes.

Reverse transcription

[257] As used herein, the term "reverse transcription" indicates the capability of an enzyme to synthesize a DNA strand (that is, complementary DNA or cDNA) using RNA as a template. In some embodiments, the reverse transcription can be "error-prone reverse transcription," which refers to the properties of certain reverse transcriptase enzymes which are error-prone in their DNA polymerization activity.

Protein, peptide, and polypeptide

[258] The terms "protein," "peptide," and "polypeptide" are used interchangeably herein, and refer to a polymer of amino acid residues linked together by peptide (amide) bonds. The terms refer to a protein, peptide, or polypeptide of any size, structure, or function. Typically, a protein, peptide, or polypeptide will be at least three amino acids long. A protein, peptide, or polypeptide may refer to an individual protein or a collection of proteins. One or more of the amino acids in a protein, peptide, or polypeptide may be modified, for example, by the addition of a chemical entity such as a carbohydrate group, a hydroxyl group, a phosphate group, a farnesyl group, an isofarnesyl group, a fatty acid group, a linker for conjugation, functionalization, or other modification, etc. A protein, peptide, or polypeptide may also be a single molecule or may be a multi-molecular complex. A protein, peptide, or polypeptide may be just a fragment of a naturally occurring protein or peptide. A protein, peptide, or polypeptide may be naturally occurring, recombinant, or synthetic, or any combination thereof. Any of the proteins provided herein may be produced by any method known in the art. For example, the proteins provided herein may be produced via recombinant protein expression and purification, which is especially suited for fusion proteins comprising a peptide linker. Methods for recombinant protein expression and purification are well known, and include those described by Green and Sambrook, Molecular Cloning: A Laboratory Manual (^{4}th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)), the entire contents of which are incorporated herein by reference.

Silent mutation

[259] As used herein, the term "silent mutation" refers to a mutation in a nucleic acid molecule that does not have an effect on the phenotype of the nucleic acid molecule, or the protein it produces if it encodes a protein. Silent mutations can be present in coding regions of a nucleic acid (*i.e.*, segments of a gene that encode for a protein), or they can be present in non-coding regions of a nucleic acid. A silent mutation in a nucleic acid sequence, e.g., in a target DNA sequence or in a DNA synthesis template sequence to be installed in the target sequence, may be a nucleotide alteration that does not result in expression or function of the amino acid sequence encoded by the nucleic acid sequence, or other functional features of the target nucleic acid sequence. When silent mutations are present in a coding region, they may be synonymous mutations. Synonymous mutations refer to substitutions of one base for another in a gene such that the corresponding amino acid residue of the protein produced by the gene is not modified. This is due to the redundancy of the genetic code, allowing for multiple different codons to encode for the same amino acid in a particular organism. When a silent mutation is in a non-coding region or a junction of a coding region and a non-coding region (e.g., an intron/exon junction), it may be in a region that does not impact any biological properties of the nucleic acid molecule (*e.g.*, splicing, gene regulation, RNA lifetime, *etc.*). Silent mutations may be useful, for example, for increasing the length of an edit made to a target nucleotide sequence using prime editing to evade correction of the edit by the MMR pathway as described herein. In certain embodiments, the number of silent mutations installed may be one, or two, or three, or four, or five, or six, or seven, or eight, or nine, or ten, or more. In certain other embodiments involving at least two silent mutations, the silent mutations may be installed within one, or two, or three, or four, or five, or six, or seven, or eight, or nine, or ten, or 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides from the intended edit site.

Spacer sequence

[260] As used herein, the term "spacer sequence" in connection with a guide RNA or a PEgRNA refers to the portion of the guide RNA or PEgRNA of about 20 nucleotides (e.g., 16, 17, 18, 19, 20, 21, 22, 23 or 24 nucleotides) which contains a nucleotide sequence that shares the same sequence as the protospacer sequence in the target DNA sequence. The spacer sequence anneals to the complement of the protospacer sequence to form a ssRNA/ssDNA hybrid structure at the target site and a corresponding R loop ssDNA structure of the endogenous DNA strand.

Target site

[261] The term "target site" refers to a sequence within a nucleic acid molecule to be edited by a prime editor (PE) disclosed herein. The target site may refer to the endogenous sequence within the nucleic acid molecule to be edited, e.g., endogenous genomic sequence of a target genome, which is identical to the sequence of the DNA synthesis template except for the one or more nucleotide edits to be installed present on the DNA synthesis template (and except that the DNA synthesis template contains Uracil instead of Thymine), or the corresponding endogenous sequence on the non-target strand that is complementary to the DNA synthesis template except for one or more mismatches at the position of the one or more nucleotide edits to be installed present on the DNA synthesis template. The target site may also further refer to the sequence within a nucleic acid molecule to which a complex of the prime editor (PE) and gRNA binds.

Variant

[262] As used herein the term "variant" should be taken to mean the exhibition of qualities that have a pattern that deviates from what occurs in nature, e.g., a variant Cas9 is a Cas9 comprising one or more changes in amino acid residues as compared to a wild type Cas9 amino acid sequence. The term "variant" encompasses homologous proteins having at least 75%, or at least 80%, or at least 85%, or at least 90%, or at least 95%, or at least 99% percent identity with a reference sequence and having the same or substantially the same functional activity or activities as the reference sequence. The term also encompasses mutants, truncations, or domains of a reference sequence, and which display the same or substantially the same functional activity or activities as the reference sequence.

Vector

[263] The term "vector," as used herein, refers to a nucleic acid that can be modified to encode a gene of interest and that is able to enter into a host cell, mutate and replicate within the host cell, and then transfer a replicated form of the vector into another host cell. Exemplary suitable vectors include viral vectors, such as retroviral vectors or bacteriophages and filamentous phage, and conjugative plasmids. Additional suitable vectors will be apparent to those of skill in the art based on the instant disclosure.

DETAILED DESCRIPTION OF CERTAIN EMBODIMENTS

[264] The present disclosure provides compositions and methods for prime editing with improved editing efficiency and/or reduced indel formation by inhibiting the DNA mismatch repair pathway while conducting prime editing of a target site. The inventors have surprisingly found that the editing efficiency of prime editing may be significantly increased (e.g., 2-fold increase, 3-fold increase, 4-fold increase, 5-fold increase, 6-fold increase, 7-fold increase, 8-fold increase, 9-fold increase, 10-fold increase, 11-fold increase, 12-fold increase,13-fold increase, 14-fold increase, 15-fold increase, 16-fold increase, 17-fold increase, 18-fold increase, 19-fold increase, 20-fold increase, 21-fold increase, 22-fold increase, 23-fold increase, 24-fold increase, 26-fold increase, 27-fold increase,28-fold increase, 29-fold increase, 30-fold increase, 31-fold increase, 32-fold increase, 33-fold increase, 34-fold increase, 35-fold increase, 36-fold increase, 37-fold increase, 38-fold increase, 39-fold increase, 40-fold increase ,41-fold increase, 42-fold increase, 43-fold increase, 44-fold increase, 45-fold increase, 46-fold increase, 47-fold increase, 48-fold increase, 49-fold increase, 50-fold increase, 51-fold increase, 52-fold increase, 53-fold increase, 54-fold increase, 55-fold increase, 56-fold increase, 57-fold increase, 58-fold increase, 59-fold increase, 60-fold increase, 61-fold increase, 62-fold increase, 63-fold increase, 64-fold increase, 65-fold increase, 66-fold increase, 67-fold increase, 68-fold increase, 69-fold increase, 70-fold increase, 71-fold increase, 72-fold increase, 73-fold increase, 74-fold increase, 75-fold increase, 76-fold increase, 77-fold increase, 78-fold increase, 79-fold increase, 80-fold increase, 81-fold increase, 82-fold increase, 83-fold increase, 84-fold increase, 85-fold increase, 86-fold increase, 87-fold increase, 88-fold increase, 89-fold increase, 90-fold increase, 91-fold increase, 92-fold increase, 93-fold increase, 94-fold increase, 95-fold increase, 96-fold increase, 97-fold increase, 98-fold increase, 99-fold increase, 100-fold increase or more) when one or more functions of the DNA mismatch repair (MMR) system are inhibited, blocked, or otherwise inactivated during prime editing. In addition, the inventors have surprisingly found that the frequency of indel formation resulting from prime editing may be significantly decreased (e.g., 2-fold decrease, 3-fold decrease, 4-fold decrease, 5-fold decrease, 6-fold decrease, 7-fold decrease, 8-fold decrease, 9-fold decrease, or 10-fold decrease or lower) when one or more functions of the DNA mismatch repair (MMR) system are inhibited, blocked, or otherwise inactivated during prime editing.

[265] The disclosure relates to the surprising finding that the efficiency and/or specificity of prime editing is impacted by a cell's own DNA mismatch repair (MMR) DNA repair pathway. As described herein (e.g., in Example 1), the inventors developed a novel genetic screening method—referred to in one embodiment as "pooled CRISPRi screen for prime editing outcomes"—which led to the identification of various genetic determinates, including MMR, as affecting the efficiency and/or specificity of prime editing. Accordingly, in one aspect, the present disclosure provides novel prime editing systems comprising a means for inhibiting and/or evade the effects of MMR, thereby increasing the efficiency and/or specificity of prime editing. In one embodiment, the disclosure provides a prime editing system that comprises an MMR-inhibiting protein, such as, but not limited to, a dominant negative MMR protein, such as a dominant negative MLH1 protein (i.e., "MLH1dn"). In another embodiment, the prime editing system comprises the installation of one or more silent mutations nearby an intended edit, thereby allowing the intended edit from evading MMR recognition, even in the absence of an MMR-inhibiting protein, such as an MLH1dn. In another aspect, the disclosure provides a novel genetic screen for identifying genetic determinants, such as MMR, that impact the efficiency and/or specificity of prime editing. In still further aspects, the disclosure provides nucleic acid constructs encoding the improved prime editing systems described herein. The disclosure in other aspects also provides vectors (e.g., AAV or lentivirus vectors) comprising nucleic acids encoding the improved prime editing system described herein. In still other aspects, the disclosure provides cells comprising the improved prime editing systems described herein. The disclosure also provides in other aspects the components of the genetic screens, including nucleic acid and/or vector constructs, guide RNA, pegRNAs, cells (e.g., CRISPRi cells), and other reagents and/or materials for conducting the herein disclosed genetic screens. In still other aspects, the disclosure provides compositions and kits, e.g., pharmaceutical compositions, comprising the improved prime editing system described herein and which are capable of being administered to a cell, tissue, or organism by any suitable means, such as by gene therapy, mRNA delivery, virus-like particle delivery, or ribonucleoprotein (RNP) delivery. In yet another aspect, the present disclosure provides methods of using the improved prime editing system to install one or more edits in a target nucleic acid molecule, e.g., a genomic locus. In still another aspect, the present disclosure provides methods of treating a disease or disorder using the improved prime editing system to correct or otherwise repair one or more genetic changes (e.g., a

single polymorphism) in a target nucleic acid molecule, e.g., a genomic locus comprising one or more disease-causing mutations.

[266] In one embodiment, the MLH1 protein is inhibited, blocked, or otherwise inactivated. In other embodiments, other proteins of the MMR system are inhibited, blocked, or otherwise inactivated, including, but not limited to, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. The inhibition may involve inhibiting the protein with an inhibitor (e.g., antibody or small molecule inhibitor or a dominant negative variant of the protein which disrupts, blocks, or otherwise inactivates the function of the protein, e.g., a dominant negative form of MLH1). The inhibition may also involve any other suitable means, such as by protein degradation (e.g., PROTAC-based degradation of MLH1), transcript-level inhibition (e.g., siRNA transcript degradation / gene silencing or microRNA-based inhibition of translation of the MLH1 transcript), or at the genetic level (i.e., installing a mutation in the MLH1 gene (or regulatory regions) which inactivates or reduces the expression of the MLH1 gene, or which installs a mutation which inactivates, blocks, or minimizes that activity of the encoded MLH1 product). In addition, the disclosure contemplates that the prime editor (e.g., delivered as a fusion protein comprising a napDNAbp and a polymerase, such as a Cas9 nickase fused to a reverse transcriptase) may be administered together with any inhibitor of the DNA mismatch repair pathway.

[267] Accordingly, the present disclosure provides a method for editing a nucleic acid molecule by prime editing that involves contacting a nucleic acid molecule with a prime editor, a pegRNA, and an inhibitor of the DNA mismatch repair pathway, thereby installing one or more modifications to the nucleic acid molecule at a target site with increased editing efficiency and/or lower indel formation. The present disclosure further provides polynucleotides for editing a DNA target site by prime editing comprising a nucleic acid sequence encoding a napDNAbp, a polymerase, and an inhibitor of the DNA mismatch repair pathway, wherein the napDNAbp and polymerase is capable in the presence of a pegRNA of installing one or more modifications in the DNA target site with increased editing efficiency and/or lower indel formation. The disclosure further provides, vectors, cells, and kits comprising the compositions and polynucleotides of the disclosure, as well as methods of making such vectors, cells, and kits, as well as methods for delivery such compositions, polynucleotides, vectors, cells and kits to cells

in vitro, *ex vivo* (e.g., during cell-based therapy which modify cells outside of the body), and *in vivo*.

MMR pathway

[268] As noted above, the present disclosure relates to the observation that the efficiency and/or specificity of prime editing is impacted by a cell's own DNA mismatch repair (MMR) DNA repair pathway. DNA mismatch repair (MMR) is a highly conserved biological pathway that plays a key role in maintaining genomic stability (*e.g.*, see FIG. 8A and 8B). *Escherichia coli* MutS and MutL and their eukaryotic homologs, MutSα and MutLα, respectively, are key players in MMR-associated genome maintenance. In various aspects, the disclosure contemplates any suitable means by which to inhibit, block, or otherwise inactivate the DNA mismatch repair (MMR) system, including, but not limited to inactivating one or more critical proteins of the MMR system at the genetic level, e.g., by introducing one or more mutations in the gene(s) encoding a protein of the MMR system. Such proteins include, but are not limited to MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. The nucleotide and amino acid sequences of such naturally occurring proteins and variants thereof are known in the art. Exemplary sequences are provided herein. The present disclosure embraces using any inhibitor, blocking agent, knockdown strategy, or other means of inactivating any known protein involved in MMR ("MMR protein"), including any wild type or naturally occurring variant of such MMR protein, and any engineered variant (including single or multiple amino acid substitutions, deletions, insertions, rearrangements, or fusions) of such MMR protein, so long as the inhibiting, blocking, or otherwise inactivation of one or more of said MMR proteins or variants thereof result in the inhibition, blockage, or inactivation of the MMR pathway. The inhibiting, blocking, or inactivation of any one or more MMR proteins or variants may be by any suitable means applied at the genetic level (e.g., in the gene encoding the one or more MMR proteins, such as introducing a mutation that inactivates the MMR protein or variant thereof), transcriptional level (e.g., by transcript knockdown), translational level (e.g., by blocking translation of one or more MMR proteins from their cognate transcripts), or at the protein level (e.g., administering of an inhibitor (e.g., small molecule, antibody, dominant negative protein variant) or by targeted protein degradation (e.g., PROTAC-based degradation).

[269] In one aspect, the present disclosure provides an improved method of prime editing comprising additionally inhibiting the DNA mismatch repair (MMR) system during prime editing by inhibiting, blocking, or otherwise inactivating MLH1 or a variant thereof.

[270] Without being bound by theory, MLH1 is a key MMR protein that heterodimerizes with PMS2 to form MutL alpha, a component of the post-replicative DNA mismatch repair system (MMR). DNA repair is initiated by MutS alpha (MSH2-MSH6) or MutS beta (MSH2-MSH3) binding to a dsDNA mismatch, then MutL alpha is recruited to the heteroduplex. Assembly of the MutL-MutS-heteroduplex ternary complex in presence of RFC and PCNA is sufficient to activate endonuclease activity of PMS2. It introduces single-strand breaks near the mismatch and thus generates new entry points for the exonuclease EXO1 to degrade the strand containing the mismatch. DNA methylation would prevent cleavage and therefore assure that only the newly mutated DNA strand is going to be corrected. MutL alpha (MLH1-PMS2) interacts physically with the clamp loader subunits of DNA polymerase III, suggesting that it may play a role to recruit the DNA polymerase III to the site of the MMR. Also implicated in DNA damage signaling, a process which induces cell cycle arrest and can lead to apoptosis in case of major DNA damages. MLH1 also heterodimerizes with MLH3 to form MutL gamma which plays a role in meiosis.

[271] The "canonical" human MLH1 amino acid sequence is represented by SEQ ID NO: 204.

[272] MLH1 also may include other human isoforms, including P40692-2 (SEQ ID NO: 205), which differs from the canonical sequence in that residues 1-241 of the canonical sequence are missing.

[273] MLH1 also may include a third known isoform known as P40692-3 (SEQ ID NO: 207), which differs from the canonical sequence in that residues 1-101 (of MSFVAGVIRR...ASISTYGFRG (SEQ ID NO: 206)) are replaced with MAF.

MMR inhibitors and methods of MMR inhibition

[274] The present disclosure provides a method for editing a nucleic acid molecule by prime editing that involves contacting a nucleic acid molecule with a prime editor, a pegRNA, and an inhibitor of the DNA mismatch repair pathway, thereby installing one or more modifications to the nucleic acid molecule at a target site with increased editing efficiency and/or lower indel formation. Thus, the present disclosure contemplates any suitable means to inhibit MMR. In one embodiment, the disclosure embraces administering an effective amount of an inhibitor of the

MMR pathway. In various embodiments, the MMR pathway may be inhibited by inhibiting, blocking, or inactivating any one or more MMR proteins or variants at the genetic level (e.g., in the gene encoding the one or more MMR proteins, such as introducing a mutation that inactivates the MMR protein or variant thereof), transcriptional level (e.g., by transcript knockdown), translational level (e.g., by blocking translation of one or more MMR proteins from their cognate transcripts), or at the protein level (e.g., application of an inhibitor (e.g., small molecule, antibody, dominant negative protein partner) or by targeted protein degradation (e.g., PROTAC-based degradation). The present disclosure also contemplates methods of prime editing which are designed to install modifications to a nucleic acid molecule that evade correction by the MMR pathway, without the need to provide an MMR inhibitor.

[275] The inventors developed prime editing which enables the insertion, deletion, or replacement of genomic DNA sequences without requiring error-prone double-strand DNA breaks. The present disclosure now provides an improved method of prime editing involving the blocking, inhibiting, or inactivation of the MMR pathway (e.g., by inhibiting, blocking, or inactivating an MMR pathway protein, including MLH1) during prime editing, whereby doing so surprisingly results in increased editing efficiency and reduced indel formation. As used herein, "during" prime editing can embrace any suitable sequence of events, such that the prime editing step can be applied before, at the same time, or after the step of blocking, inhibiting, or inactivating the MMR pathway (e.g., by targeting the inhibition of MLH1).

[276] Prime editing uses an engineered Cas9 nickase–reverse transcriptase fusion protein (e.g., PE1 or PE2) paired with an engineered prime editing guide RNA (pegRNA) that both directs Cas9 to the target genomic site and encodes the information for installing the desired edit. Prime editing proceeds through a multi-step editing process: 1) the Cas9 domain binds and nicks the target genomic DNA site, which is specified by the pegRNA's spacer sequence; 2) the reverse transcriptase domain uses the nicked genomic DNA as a primer to initiate the synthesis of an edited DNA strand using an engineered extension on the pegRNA as a template for reverse transcription–this generates a single-stranded 3′ flap containing the edited DNA sequence; 3) cellular DNA repair resolves the 3′ flap intermediate by the displacement of a 5' flap species that occurs via invasion by the edited 3′ flap, excision of the 5′ flap containing the original DNA sequence, and ligation of the new 3′ flap to incorporate the edited DNA strand, forming a heteroduplex of one edited and one unedited strand; and 4) cellular DNA repair replaces the

unedited strand within the heteroduplex using the edited strand as a template for repair, completing the editing process.

[277] Efficient incorporation of the desired edit requires that the newly synthesized 3′ flap contains a portion of sequence that is homologous to the genomic DNA site. This homology enables the edited 3′ flap to compete with the endogenous DNA strand (the corresponding 5' flap) for incorporation into the DNA duplex. Because the edited 3' flap will contain less sequence homology than the endogenous 5′ flap, the competition is expected to favor the 5′ flap strand. Thus, a potential limiting factor in the efficiency of prime editing may be the failure of the 3′ flap, which contains the edit, to effectively invade and displace the 5′ flap strand. Moreover, successful 3′ flap invasion and removal of the 5′ flap only incorporates the edit on one strand of the double-stranded DNA genome. Permanent installation of the edit requires cellular DNA repair to replace the unedited complementary DNA strand using the edited strand as a template. While the cell can be made to favor replacement of the unedited strand over the edited strand (step 4 above) by the introduction of a nick in the unedited strand adjacent to the edit using a secondary sgRNA (the PE3 system), this process still relies on a second stage of DNA repair.

[278] This disclosure describes a modified approach to prime editing that comprises additionally inhibiting, blocking, or otherwise inactivating the DNA mismatch repair (MMR) system. In some embodiments, an MMR inhibitor is provided to the target nucleic acid along with other components of a prime editing system, for example, an exogenous MMR inhibitor such as an siRNA can be provided to a cell comprising the target nucleic acid. In some embodiments, a prime editing system component, *e.g.*, a pegRNA, is designed to install modifications in the target nucleic acid which evade the MMR system, without the need to provide an inhibitor. In certain embodiments, the DNA mismatch repair (MMR) system can be inhibited, blocked, or otherwise inactivating one or more proteins of the MMR system, including, but not limited to MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. The disclosure contemplates any suitable means by which to inhibit, block, or otherwise inactivate the DNA mismatch repair (MMR) system, including, but not limited to inactivating one or more critical proteins of the MMR system at the genetic level, e.g., by introducing one or more mutations in the genes encoding a protein of the MMR system,

e.g., MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

[279] Thus, in one aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating the DNA mismatch repair (MMR) system.

[280] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating a protein of the MMR system, e.g., MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

[281] In one aspect, the present disclosure provides an improved method of prime editing comprising additionally inhibiting the DNA mismatch repair (MMR) system during prime editing by inhibiting, blocking, or otherwise inactivating MLH1 or a variant thereof. Without being bound by theory, MLH1 is a key MMR protein that heterodimerizes with PMS2 to form MutL alpha, a component of the post-replicative DNA mismatch repair system (MMR). DNA repair is initiated by MutS alpha (MSH2-MSH6) or MutS beta (MSH2-MSH3) binding to a dsDNA mismatch, then MutL alpha is recruited to the heteroduplex. Assembly of the MutL-MutS-heteroduplex ternary complex in presence of RFC and PCNA is sufficient to activate endonuclease activity of PMS2. It introduces single-strand breaks near the mismatch and thus generates new entry points for the exonuclease EXO1 to degrade the strand containing the mismatch. DNA methylation would prevent cleavage and therefore assure that only the newly mutated DNA strand is going to be corrected. MutL alpha (MLH1-PMS2) interacts physically with the clamp loader subunits of DNA polymerase III, suggesting that it may play a role to recruit the DNA polymerase III to the site of the MMR. Also implicated in DNA damage signaling, a process which induces cell cycle arrest and can lead to apoptosis in case of major DNA damages. MLH1 also heterodimerizes with MLH3 to form MutL gamma which plays a role in meiosis. The "canonical" human MLH1 amino acid sequence is represented by SEQ ID NO: 204

[282] MLH1 also may include other human isoforms, including P40692-2 (SEQ ID NO: 205), which differs from the canonical sequence in that residues 1-241 of the canonical sequence are missing.

[283] MLH1 also may include a third known isoform known as P40692-3 (SEQ ID NO: 207), which differs from the canonical sequence in that residues 1-101 (of MSFVAGVIRR…ASISTYGFRG (SEQ ID NO: 206)) are replaced with MAF.

[284] The disclosure contemplates that any of the following MLH1 proteins may be inhibited by an inhibitor, or otherwise blocked or inactivated in order to inhibit the MMR pathway during prime editing. In addition, such exemplary proteins may also be used to engineer or otherwise make a dominant negative variant that may be used as a type of inhibitor when administered in an effective amount which blocks, inactivates, or inhibits the MMR. Without being bound by theory, it is believed that MLH1 dominant negative mutants can saturate binding of MutS. Exemplary MLH1 proteins include the following amino acid sequences, or amino acid sequences having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to 100% sequence identity with any of the following sequences:

Description	Sequence	SEQ ID NO:
MLH1 *Homo sapiens* SwissProt Accession No. P40692 Wild type	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGG LKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAH VTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKAL KNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFG NAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIE TVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIES KLLGSNSSRMYFTQTLLPGLAGPSGEMVKSTTSLTSSSTSGSSDKVYAHQMVRT DSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELPAPAEVA AKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAAC TPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLY LLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGP KEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRL ATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWT VEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVFERC	204
MLH1 *Mus musculus* SwissProt Accession No. Q9JK91 Wild type	MAFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTNIQVVVKEG GLKLIQIQDNGTGIRKEDLDIVCERFTTSKLQTFEDLASISTYGFRGEALASISHV AHVTITTKTADGKCAYRASYSDGKLQAPPKPCAGNQGTLITVEDLFYNIITRRK ALKNPSEEYGKILEVVGRYSIHNSGISFSVKKQGETVSDVRTLPNATTVDNIRSIF GNAVSRELIEVGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESAALRK AIETVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILQRVQQH IESKLLGSNSSRMYFTQTLLPGLAGPSGEAARPTTGVASSSTSGSGDKVYAYQM VRTDSREQKLDAFLQPVSSLGPSQPQDPAPVRGARTEGSPERATREDEEMLALP APAEAAAESENLERESLMETSDAAQKAAPTSSPGSSRKRHREDSDVEMVENAS GKEMTAACYPRRRIINLTSVLSLQEEISERCHETLREMLRNHSFVGCVNPQWAL	219

	AQHQTKLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPES GWTEDDGPKEGLAEYIVEFLKKKAEMLADYFSVEIDEEGNLIGLPLLIDSYVPPL EGLPIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYILEESTLSGQQSDMPG STSKPWKWTVEHIIYKAFRSHLLPPKHFTEDGNVLQLANLPDLYKVFERC	
MLH1 *Rattus norvegicus* SwissProt Accession No. P97679 Wild type	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMTENCLDAKSTNIQVIVREG GLKLIQIQDNGTGIRKEDLDIVCERFTTSKLQTFEDLAMISTYGFRGEALASISHV AHVTITTKTADGKCAYRASYSDGKLQAPPKPCAGNQGTLITVEDLFYNIITRKK ALKNPSEEYGKILEVVGRYSIHNSGISFSVKKQGETVSDVRTLPNATTVDNIRSIF GNAVSRELIEVGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESAALKK AIEAVYAAYLPKNTHPFLYLILEISPQNVDVNVHPTKHEVHFLHEESILERVQQHI ESKLLGSNSSRMYFTQTLLPGLAGPSGEAVKSTTGIASSSTSGSGDKVHAYQMV RTDSRDQKLDAFMQPVSRRLPSQPQDPVPGNRTEGSPEKAMQKDQEISELPAPM EAAADSASLERESVIGASEVVAPQRHPSSPGSSRKRHPEDSDVEMMENDSRKEM TAACYPRRRIINLTSVLSLQEEINDRGHETLREMLRNHTFVGCVNPQWALAQHQ TKLYLLNTTKLSEELFYQILIYDFANFGVLRLPEPAPLFDFAMLALDSPESGWTE EDGPKEGLAEYIVEFLKKKAKMLADYFSVEIDEEGNLIGLPLLIDSYVPPLEGLPI FILRLATEVNWDEEECFESLSKECAVFYSIRKQYILEESALSGQQSDMPGSPSKP WKWTVEHIIYKAFRSHLLPPKHFTEDGNVLQLANLPDLCKVFERC	220
MLH1 *Bos taurus* SwissProt Accession No. F1MPGO Wild type	MSLVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVVVKEG GLKLIQIQDNGTGIRKEDLEIVCERFTTSKLQSFEDLAHISTYGFRGEALASISHV AHVTITTKTADGKCAYRAHYSDGKLKAPPKPCAGNQGTQITVEDLFYNISTRRK ALKNPSEEYGKILEVVGRYAVHNSGIGFSVKKQGETVADVRTLPNATTVDNIRS IFGNAVSRELIEVECEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESASLRK AIETVYAAYLPKSTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEDSILERLQQHI ESRLLGSNASRTYFTQTLLPGLPGPSGEAVKSTASVTSSSTAGSGDRVYAHQMV RTDCREQKLDAFLQPVSKALSSQPQAVVPEHRTDASSSGTRQQDEEMLELPAPA AVAAKSQALEDDATMRAADLAEKRGPSSSPENPRKRPREDSDVEMVEDASRKE MTAACTPRRRIINLTSVLSLQEEINERGHETLREMLHNHSFVGCVNPQWALAQH QTKLYLLNTTRLSEELFYQILVYDFANFGVLRLSEPAPLFDLAMLALDSPESGWT EEDGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLVGLPLLIDNYVPPLEGL PIFILRLATEVNWDEEKECFESLSKECAMFYSIRKQYVSAESTLSGQQSEVPGST ANPWKWTVEHVIYKAFRSHLLPPKHFTEDGNILQLANLPDLYKVFERC	221

[285] The methods and compositions described herein utilize MLH1 mutants or truncated variants. In some embodiments, the mutants and truncated variants of the human MLH1 wild-type protein are utilized.

[286] In one aspect, a truncated variant of human MLH1 is provided by this disclosure. In some embodiments, amino acids 754-756 of the wild-type human MLH1 protein are truncated (Δ754-756, hereinafter referred to as MLH1dn). In some embodiments, a truncated variant of human MLH1 comprising only the N-terminal domain (amino acids 1-335) is provided (hereinafter referred to as $MLH1dn^{NTD}$. In various embodiments, the following MLH1 variants are provided in this disclosure:

Description	Sequence	SEQ ID NO:
MLH1 E34A	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKAMIENCLDAKSTSIQVIVKEGG LKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAH VTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKAL KNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFG NAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIE	222

	TVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIES KLLGSNSSRMYFTQTLLPGLAGPSGEMVKSTTSLTSSSTSGSSDKVYAHQMVRT DSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELPAPAEVA AKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAAC TPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLY LLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGP KEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRL ATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWT VEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVFERC	
MLH1 Δ756	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGG LKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAH VTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKAL KNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFG NAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIE TVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIES KLLGSNSSRMYFTQTLLPGLAGPSGEMVKSTTSLTSSSTSGSSDKVYAHQMVRT DSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELPAPAEVA AKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAAC TPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLY LLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGP KEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRL ATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWT VEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVFER{-}	208
MLH1 Δ754-756	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGG LKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAH VTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKAL KNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFG NAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIE TVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIES KLLGSNSSRMYFTQTLLPGLAGPSGEMVKSTTSLTSSSTSGSSDKVYAHQMVRT DSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELPAPAEVA AKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAAC TPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLY LLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGP KEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRL ATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWT VEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVF{- - -}	209
MLH1 E34A Δ754-756	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIK<u>A</u>MIENCLDAKSTSIQVIVKEGG LKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAH VTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKAL KNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFG NAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIE TVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIES KLLGSNSSRMYFTQTLLPGLAGPSGEMVKSTTSLTSSSTSGSSDKVYAHQMVRT DSREQKLDAFLQPLSKPLSSQPQAIVTEDKTDISSGRARQQDEEMLELPAPAEVA AKNQSLEGDTTKGTSEMSEKRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAAC TPRRRIINLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLY LLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGP KEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRL ATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWT VEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVF{- - -}	210
MLH1 1-335	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGG LKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAH VTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKAL KNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFG	211

	NAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIE TVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIES KLL	
MLH1 1-335 E34A	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIK**A**MIENCLDAKSTSIQVIVKEGG LKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAH VTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKAL KNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFG NAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIE TVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIES KLL	212
MLH1 1-335 NLS^{SV40}	MSFVAGVIRRLDETVVNRIAAGEVIQRPANAIKEMIENCLDAKSTSIQVIVKEGG LKLIQIQDNGTGIRKEDLDIVCERFTTSKLQSFEDLASISTYGFRGEALASISHVAH VTITTKTADGKCAYRASYSDGKLKAPPKPCAGNQGTQITVEDLFYNIATRRKAL KNPSEEYGKILEVVGRYSVHNAGISFSVKKQGETVADVRTLPNASTVDNIRSIFG NAVSRELIEIGCEDKTLAFKMNGYISNANYSVKKCIFLLFINHRLVESTSLRKAIE TVYAAYLPKNTHPFLYLSLEISPQNVDVNVHPTKHEVHFLHEESILERVQQHIES KLL**PKKKRKV**	213
MLH1 501-756	INLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTT KLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGPKEGLA EYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVN WDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIV YKALRSHILPPKHFTEDGNILQLANLPDLYKVFERC	215
MLH1 501-753	INLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTT KLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGPKEGLA EYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVN WDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIV YKALRSHILPPKHFTEDGNILQLANLPDLYKVF[- - -]	216
MLH1 461-756	KRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINE QGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDF ANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEML ADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFESLSKE CAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFT EDGNILQLANLPDLYKVFERC	217
MLH1 461-753	KRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSVLSLQEEINE QGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEELFYQILIYDF ANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEFLKKKAEML ADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEKECFESLSKE CAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALRSHILPPKHFT EDGNILQLANLPDLYKVF[- - -]	218
NLS^{SV40} MLH1 501-753	**PKKKRKV**INLTSVLSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQT KLYLLNTTKLSEELFYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEE DGPKEGLAEYIVEFLKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFI LRLATEVNWDEEKECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSW KWTVEHIVYKALRSHILPPKHFTEDGNILQLANLPDLYKVF[- - -]	223
NLS^{SV40} MLH1 461-753	**PKKKRKV**KRGPTSSNPRKRHREDSDVEMVEDDSRKEMTAACTPRRRIINLTSV LSLQEEINEQGHEVLREMLHNHSFVGCVNPQWALAQHQTKLYLLNTTKLSEEL FYQILIYDFANFGVLRLSEPAPLFDLAMLALDSPESGWTEEDGPKEGLAEYIVEF LKKKAEMLADYFSLEIDEEGNLIGLPLLIDNYVPPLEGLPIFILRLATEVNWDEEK ECFESLSKECAMFYSIRKQYISEESTLSGQQSEVPGSIPNSWKWTVEHIVYKALR SHILPPKHFTEDGNILQLANLPDLYKVF[- - -]	224

[287] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PMS2 (or MutL alpha) or variant thereof.

[288] In yet another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PMS1 (or MutL beta) or variant thereof.

[289] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MLH3 (or MutL gamma) or variant thereof.

[290] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MutS alpha (MSH2-MSH6) or variant thereof.

[291] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MSH2 or variant thereof.

[292] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MSH6 or variant thereof.

[293] In yet another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PCNA or variant thereof.

[294] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating RFC or variant thereof.

[295] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating EXO1or variant thereof.

[296] In yet another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating POLδ or variant thereof.

[297] Exemplary amino acid sequences for these MMR proteins (PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA) are as follows:

Description	Sequence	SEQ ID NO:
PMS2 *Homo sapiens* SwissProt Accession No. P54278 Wild type	MERAESSSTEPAKAIKPIDRKSVHQICSGQVVLSLSTAVKELVENSLDAGATNIDL KLKDYGVDLIEVSDNGCGVEEENFEGLTLKHHTSKIQEFADLTQVETFGFRGEAL SSLCALSDVTISTCHASAKVGTRLMFDHNGKIIQKTPYPRPRGTTVSVQQLFSTLP VRHKEFQRNIKKEYAKMVQVLHAYCIISAGIRVSCTNQLGQGKRQPVVCTGGSP SIKENIGSVFGQKQLQSLIPFVQLPPSDSVCEEYGLSCSDALHNLFYISGFISQCTHG VGRSSTDRQFFFINRRPCDPAKVCRLVNEVYHMYNRHQYPFVVLNISVDSECVDI NVTPDKRQILLQEEKLLLAVLKTSLIGMFDSDVNKLNVSQQPLLDVEGNLIKMH AADLEKPMVEKQDQSPSLRTGEEKKDVSISRLREAFSLRHTTENKPHSPKTPEPR RSPLGQKRGMLSSSTSGAISDKGVLRPQKEAVSSSHGPSDPTDRAEVEKDSGHGS TSVDSEGFSIPDTGSHCSSEYAASSPGDRGSQEHVDSQEKAPKTDDSFSDVDCHS NQEDTGCKFRVLPQPTNLATPNTKRFKKEEILSSSDICQKLVNTQDMSASQVDVA VKINKKVVPLDFSMSSLAKRIKQLHHEAQQSEGEQNYRKFRAKICPGENQAAED ELRKEISKTMFAEMEIIGQFNLGFIITKLNEDIFIVDQHATDEKYNFEMLQQHTVL QGQRLIAPQTLNLTAVNEAVLIENLEIFRKNGFDFVIDENAPVTERAKLISLPTSKN WTFGPQDVDELIFMLSDSPGVMCRPSRVKQMFASRACRKSVMIGTALNTSEMK KLITHMGEMDHPWNCPHGRPTMRHIANLGVISQN	225
PMS1 *Homo sapiens* SwissProt Accession No. P54277 Wild type	MKQLPAATVRLLSSSQIITSVVSVVKELIENSLDAGATSVDVKLENYGFDKIEVRD NGEGIKAVDAPVMAMKYYTSKINSHEDLENLTTYGFRGEALGSICCIAEVLITTR TAADNFSTQYVLDGSGHILSQKPSHLGQGTTVTALRLFKNLPVRKQFYSTAKKC KDEIKKIQDLLMSFGILKPDLRIVFVHNKAVIWQKSRVSDHKMALMSVLGTAVM NNMESFQYHSEESQIYLSGFLPKCDADHSFTSLSTPERSFIFINSRPVHQKDILKLIR HHYNLKCLKESTRLYPVFFLKIDVPTADVDVNLTPDKSQVLLQNKESVLIALENL MTTCYGPLPSTNSYENNKTDVSAADIVLSKTAETDVLFNKVESSGKNYSNVDTS VIPFQNDMHNDESGKNTDDCLNHQISIGDFGYGHCSSEISNIDKNTKNAFQDISMS NVSWENSQTEYSKTCFISSVKHTQSENGNKDHIDESGENEEEAGLENSSEISADE WSRGNILKNSVGENIEPVKILVPEKSLPCKVSNNNYPIPEQMNLNEDSCNKKSNVI DNKSGKVTAYDLLSNRVIKKPMSASALFVQDHRPQFLIENPKTSLEDATLQIEEL WKTLSEEEKLKYEEKATKDLERYNSQMKRAIEQESQMSLKDGRKKIKPTSAWN LAQKHKLKTSLSNQPKLDELLQSQIEKRRSQNIKMVQIPFSMKNLKINFKKQNKV DLEEKDEPCLIHNLRFPDAWLMTSKTEVMLLNPYRVEEALLFKRLLENHKLPAEP LEKPIMLTESLFNGSHYLDVLYKMTADDQRYSGSTYLSDPRLTANGFKIKLIPGV SITENYLEIEGMANCLPFYGVADLKEILNAILNRNAKEVYECRPRKVISYLEGEAV RLSRQLPMYLSKEDIQDIIYRMKHQFGNEIKECVHGRPFFHHLTYLPETT	226
MLH3 *Homo sapiens* SwissProt Accession No. Q9UHC1 Wild type	MIKCLSVEVQAKLRSGLAISSLGQCVEELALNSIDAEAKCVAVRVNMETFQVQVI DNGFGMGSDDVEKVGNRYFTSKCHSVQDLENPRFYGFRGEALANIADMASAVE ISSKKNRTMKTFVKLFQSGKALKACEADVTRASAGTTVTVYNLFYQLPVRRKC MDPRLEFEKVRQRIEALSLMHPSISFSLRNDVSGSMVLQLPKTKDVCSRFCQIYGL GKSQKLREISFKYKEFELSGYISSEAHYNKNMQFLFVNKRLVLRTKLHKLIDFLLR KESIICKPKNGPTSRQMNSSLRHRSTPELYGIYVINVQCQFCEYDVCMEPAKTLIE FQNWDTLLFCIQEGVKMFLKQEKLFVELSGEDIKEFSEDNGFSLFDATLQKRVTS DERSNFQEACNNILDSYEMFNLQSKAVKRKTTAENVNTQSSRDSEATRKNTNDA FLYIYESGGPGHSKMTEPSLQNKDSSCSESKMLEQETIVASEAGENEKHKKSFLE HSSLENPCGTSLEMFLSPFQTPCHFEESGQDLEIWKESTTVNGMAANILKNNRIQN QPKRFKDATEVGCQPLPFATTLWGVHSAQTEKEKKKESSNCGRRNVFSYGRVKL CSTGFITHVVQNEKTKSTETEHSFKNYVRPGPTRAQETFGNRTRHSVETPDIKDL ASTLSKESGQLPNKKNCRTNISYGLENEPTATYTMFSAFQEGSKKSQTDCILSDTS	227

	PSFPWYRHVSNDSRKTDKLIGFSKPIVRKKLSLSSQLGSLEKFKRQYGKVENPLD TEVEESNGVTTNLSLQVEPDILLKDKNRLENSDVCKITTMEHSDSDSSCQPASHIL NSEKFPFSKDEDCLEQQMPSLRESPMTLKELSLFNRKPLDLEKSSESLASKLSRLK GSERETQTMGMMSRFNELPNSDSSRKDSKLCSVLTQDFCMLFNNKHEKTENGVI PTSDSATQDNSFNKNSKTHSNSNTTENCVISETPLVLPYNNSKVTGKDSDVLIRAS EQQIGSLDSPSGMLMNPVEDATGDQNGICFQSEESKARACSETEESNTCCSDWQR HFDVALGRMVYVNKMTGLSTFIAPTEDIQAACTKDLTTVAVDVVLENGSQYRC QPFRSDLVLPFLPRARAERTVMRQDNRDTVDDTVSSESLQSLFSEWDNPVFARYP EVAVDVSSGQAESLAVKIHNILYPYRFTKGMIHSMQVLQQVDNKFIACLMSTKT EENGEAGGNLLVLVDQHAAHERIRLEQLIIDSYEKQQAQGSGRKKLLSSTLIPPLE ITVTEEQRRLLWCYHKNLEDLGLEFVFPDTSDSLVLVGKVPLCFVEREANELRRG RSTVTKSIVEEFIREQLELLQTTGGIQGTLPLTVQKVLASQACHGAIKFNDGLSLQ ESCRLIEALSSCQLPFQCAHGRPSMLPLADIDHLEQEKQIKPNLTKLRKMAQAWR LFGKAECDTRQSLQQSMPPCEPP	
MSH2 *Homo sapiens* SwissProt Accession No. P43246 Wild type	MAVQPKETLQLESAAEVGFVRFFQGMPEKPTTTVRLFDRGDFYTAHGEDALLAA REVFKTQGVIKYMGPAGAKNLQSVVLSKMNFESFVKDLLLVRQYRVEVYKNRA GNKASKENDWYLAYKASPGNLSQFEDILFGNNDMSASIGVVGVKMSAVDGQRQ VGVGYVDSIQRKLGLCEFPDNDQFSNLEALLIQIGPKECVLPGGETAGDMGKLRQ IIQRGGILITERKKADFSTKDIYQDLNRLLKGKKGEQMNSAVLPEMENQVAVSSL SAVIKFLELLSDDSNFGQFELTTFDFSQYMKLDIAAVRALNLFQGSVEDTTGSQSL AALLNKCKTPQGQRLVNQWIKQPLMDKNRIEERLNLVEAFVEDAELRQTLQEDL LRRFPDLNRLAKKFQRQAANLQDCYRLYQGINQLPNVIQALEKHEGKHQKLLLA VFVTPLTDLRSDFSKFQEMIETTLDMDQVENHEFLVKPSFDPNLSELREIMNDLE KKMQSTLISAARDLGLDPGKQIKLDSSAQFGYYFRVTCKEEKVLRNNKNFSTVDI QKNGVKFTNSKLTSLNEEYTKNKTEYEEAQDAIVKEIVNISSGYVEPMQTLNDVL AQLDAVVSFAHVSNGAPVPYVRPAILEKGQGRIILKASRHACVEVQDEIAFIPND VYFEKDKQMFHIITGPNMGGKSTYIRQTGVIVLMAQIGCFVPCESAEVSIVDCILA RVGAGDSQLKGVSTFMAEMLETASILRSATKDSLIIIDELGRGTSTYDGFGLAWAI SEYIATKIGAFCMFATHFHELTALANQIPTVNNLHVTALTTEETLTMLYQVKKGV CDQSFGIHVAELANFPKHVIECAKQKALELEEFQYIGESQGYDIMEPAAKKCYLE REQGEKIIQEFLSKVKQMPFTEMSEENITIKLKQLKAEVIAKNNSFVNEIISRIKVTT	228
MSH6 *Homo sapiens* SwissProt Accession No. P52701 Wild type	MSRQSTLYSFFPKSPALSDANKASARASREGGRAAAAPGASPSPGGDAAWSEAG PGPRPLARSASPPKAKNLNGGLRRSVAPAAPTSCDFSPGDLVWAKMEGYPWWP CLVYNHPFDGTFIREKGKSVRVHVQFFDDSPTRGWVSKRLLKPYTGSKSKEAQK GGHFYSAKPEILRAMQRADEALNKDKIKRLELAVCDEPSEPEEEEEMEVGTTYV TDKSEEDNEIESEEEVQPKTQGSRRSSRQIKKRRVISDSESDIGGSDVEFKPDTKEE GSSDEISSGVGDSESEGLNSPVKVARKRKRMVTGNGSLKRKSSRKETPSATKQAT SISSETKNTLRAFSAPQNSESQAHVSGGGDDSSRPTVWYHETLEWLKEEKRRDEH RRRPDHPDFDASTLYVPEDFLNSCTPGMRKWWQIKSQNFDLVICYKVGKFYELY HMDALIGVSELGLVFMKGNWAHSGFPEIAFGRYSDSLVQKGYKVARVEQTETPE MMEARCRKMAHISKYDRVVRREICRIITKGTQTYSVLEGDPSENYSKYLLSLKEK EEDSSGHTRAYGVCFVDTSLGKFFIGQFSDDRHCSRFRTLVAHYPPVQVLFEKGN LSKETKTILKSSLSCSLQEGLIPGSQFWDASKTLRTLLEEEYFREKLSDGIGVMLPQ VLKGMTSESDSIGLTPGEKSELALSALGGCVFYLKKCLIDQELLSMANFEEYIPLD SDTVSTTRSGAIFTKAYQRMVLDAVTLNNLEIFLNGTNGSTEGTLLERVDTCHTP FGKRLLKQWLCAPLCNHYAINDRLDAIEDLMVVPDKISEVVELLKKLPDLERLLS KIHNVGSPLKSQNHPDSRAIMYEETTYSKKKIIDFLSALEGFKVMCKIIGIMEEVA DGFKSKILKQVISLQTKNPEGRFPDLTVELNRWDTAFDHEKARKTGLITPKAGFD SDYDQALADIRENEQSLLEYLEKQRNRIGCRTIVYWGIGRNRYQLEIPENFTTRNL PEEYELKSTKKGCKRYWTKTIEKKLANLINAEERRDVSLKDCMRRLFYNFDKNY KDWQSAVECIAVLDVLLCLANYSRGGDGPMCRPVILLPEDTPPFLELKGSRHPCI TKTFFGDDFIPNDILIGCEEEEQENGKAYCVLVTGPNMGGKSTLMRQAGLLAVM AQMGCYVPAEVCRLTPIDRVFTRLGASDRIMSGESTFFVELSETASILMHATAHS LVLVDELGRGTATFDGTAIANAVVKELAETIKCRTLFSTHYHSLVEDYSQNVAV	229

	RLGHMACMVENECEDPSQETITFLYKFIKGACPKSYGFNAARLANLPEEVIQKGH RKAREFEKMNQSLRLFREVCLASERSTVDAEAVHKLLTLIKEL	
PCNA *Homo sapiens* SwissProt Accession No. P12004 Wild type	MFEARLVQGSILKKVLEALKDLINEACWDISSSGVNLQSMDSSHVSLVQLTLRSE GFDTYRCDRNLAMGVNLTSMSKILKCAGNEDIITLRAEDNADTLALVFEAPNQE KVSDYEMKLMDLDVEQLGIPEQEYSCVVKMPSGEFARICRDLSHIGDAVVISCA KDGVKFSASGELGNGNIKLSQTSNVDKEEEAVTIEMNEPVQLTFALRYLNFFTKA TPLSSTVTLSMSADVPLVVEYKIADMGHLKYYLAPKIEDEEGS	230
RFC *Homo sapiens* SwissProt Accession No. P35251 Wild type	MDIRKFFGVIPSGKKLVSETVKKNEKTKSDEETLKAKKGIKEIKVNSSRKEDDFK QKQPSKKKRIIYDSDSESEETLQVKNAKKPPEKLPVSSKPGKISRQDPVTYISETDE EDDFMCKKAASKSKENGRSTNSHLGTSNMKKNEENTKTKNKPLSPIKLTPTSVL DYFGTGSVQRSNKKMVASKRKELSQNTDESGLNDEAIAKQLQLDEDAELERQL HEDEEFARTLAMLDEEPKTKKARKDTEAGETFSSVQANLSKAEKHKYPHKVKT AQVSDERKSYSPRKQSKYESSKESQQHSKSSADKIGEVSSPKASSKLAIMKRKEE SSYKEIEPVASKRKENAIKLKGETKTPKKTKSSPAKKESVSPEDSEKKRTNYQAY RSYLNREGPKALGSKEIPKGAENCLEGLIFVITGVLESIERDEAKSLIERYGGKVTG NVSKKTNYLVMGRDSGQSKSDKAAALGTKIIDEDGLLNLIRTMPGKKSKYEIAV ETEMKKESKLERTPQKNVQGKRKISPSKKESESKKSRPTSKRDSLAKTIKKETDV FWKSLDFKEQVAEETSGDSKARNLADDSSENKVENLLWVDKYKPTSLKTIIGQQ GDQSCANKLLRWLRNWQKSSSEDKKHAAKFGKFSGKDDGSSFKAALLSGPPGV GKTTTASLVCQELGYSYVELNASDTRSKSSLKAIVAESLNNTSIKGFYSNGAASS VSTKHALIMDEVDGMAGNEDRGGIQELIGLIKHTKIPIICMCNDRNHPKIRSLVHY CFDLRFQRPRVEQIKGAMMSIAFKEGLKIPPPAMNEIILGANQDIRQVLHNLSMW CARSKALTYDQAKADSHRAKKDIKMGPFDVARKVFAAGEETAHMSLVDKSDLF FHDYSIAPLFVQENYIHVKPVAAGGDMKKHLMLLSRAADSICDGDLVDSQIRSK QNWSLLPAQAIYASVLPGELMRGYMTQFPTFPSWLGKHSSTGKHDRIVQDLALH MSLRTYSSKRTVNMDYLSLLRDALVQPLTSQGVDGVQDVVALMDTYYLMKED FENIMEISSWGGKPSPFSKLDPKVKAAFTRAYNKEAHLTPYSLQAIKASRHSTSPS LDSEYNEELNEDDSQSDEKDQDAIETDAMIKKKTKSSKPSKPEKDKEPRKGKGK SSKK	231
EXO1 *Homo sapiens* SwissProt Accession No. Q9UQ84 Wild type	MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPTDR YVGFCMKFVNMLLSHGIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLR EGKVSEARECFTRSINITHAMAHKVIKAARSQGVDCLVAPYEADAQLAYLNKAG IVQAIITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFTEEKFRY MCILSGCDYLSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYIN GFIRANNTFLYQLVFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGN KDINTFEQIDDYNPDTAMPAHSRSHSWDDKTCQKSANVSSIWHRNYSPRPESGT VSDAPQLKENPSTVGVERVISTKGLNLPRKSSIVKRPRSAELSEDDLLSQYSLSFT KKTKKNSSEGNKSLSFSEVFVPDLVNGPTNKKSVSTPPRTRNKFATFLQRKNEES GAVVVPGTRSRFFCSSDSTDCVSNKVSIQPLDETAVTDKENNLHESEYGDQEGKR LVDTDVARNSSDDIPNNHIPGDHIPDKATVFTDEESYSFESSKFTRTISPPTLGTLR SCFSWSGGLGDFSRTPSPSPSTALQQFRRKSDSPTSLPENNMSDVSQLKSEESSDD ESHPLREEACSSQSQESGEFSLQSSNASKLSQCSSKDSDSEESDCNIKLLDSQSDQT SKLRLSHFSKKDTPLRNKVPGLYKSSSADSLSTTKIKPLGPARASGLSKKPASIQK RKHHNAENKPGLQIKLNELWKNFGFKKDSEKLPPCKKPLSPVRDNIQLTPEAEED IFNKPECGRVQRAIFQ	156
POLδ *Homo sapiens* SwissProt Accession No. P28340 Wild type	MDGKRRPGPGPGVPPKRARGGLWDDDDAPRPSQFEEDLALMEEMEAEHRLQEQ EEEELQSVLEGVADGQVPPSAIDPRWLRPTPPALDPQTEPLIFQQLEIDHYVGPAQ PVPGGPPPSRGSVPVLRAFGVTDEGFSVCCHIHGFAPYFYTPAPPGFGPEHMGDL QRELNLAISRDSRGGRELTGPAVLAVELCSRESMFGYHGHGPSPFLRITVALPRLV APARRLLEQGIRVAGLGTPSFAPYEANVDFEIRFMVDTDIVGCNWLELPAGKYAL RLKEKATQCQLEADVLWSDVVSHPPEGPWQRIAPLRVLSFDIECAGRKGIFPEPE RDPVIQICSLGLRWGEPEPFLRLALTLRPCAPILGAKVQSYEKEEDLLQAWSTFIRI MDPDVITGYNIQNFDLPYLISRAQTLKVQTFPFLGRVAGLCSNIRDSSFQSKQTGR RDTKVVSMVGRVQMDMLQVLLREYKLRSYTLNAVSFHFLGEQKEDVQHSIITD	232

	LQNGNDQTRRRLAVYCLKDAYLPLRLLERLMVLVNAVEMARVTGVPLSYLLSR GQQVKVVSQLLRQAMHEGLLMPVVKSEGGEDYTGATVIEPLKGYYDVPIATLD FSSLYPSIMMAHNLCYTTLLRPGTAQKLGLTEDQFIRTPTGDEFVKTSVRKGLLP QILENLLSARKRAKAELAKETDPLRRQVLDGRQLALKVSANSVYGFTGAQVGKL PCLEISQSVTGFGRQMIEKTKQLVESKYTVENGYSTSAKVVYGDTDSVMCRFGV SSVAEAMALGREAADWVSGHFPSPIRLEFEKVYFPYLLISKKRYAGLLFSSRPDA HDRMDCKGLEAVRRDNCPLVANLVTASLRRLLIDRDPEGAVAHAQDVISDLLCN RIDISQLVITKELTRAASDYAGKQAHVELAERMRKRDPGSAPSLGDRVPYVIISAA KGVAAYMKSEDPLFVLEHSLPIDTQYYLEQQLAKPLLRIFEPILGEGRAEAVLLR GDHTRCKTVLTGKVGGLLAFAKRRNCCIGCRTVLSHQGAVCEFCQPRESELYQK EVSHLNALEERFSRLWTQCQRCQGSLHEDVICTSRDCPIFYMRKKVRKDLEDQE QLLRRFGPPGPEAW	

[298] Thus, in one aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating the DNA mismatch repair (MMR) system.

[299] In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of the MMR system, e.g., an inhibitor of one or more of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, or PCNA. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an antibody, e.g., a neutralizing antibody. In still other embodiments, the inhibitor can be a dominant negative mutant of one or more of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, or PCNA, e.g., a dominant negative mutant of MLH1. In still other embodiments, the inhibitor can be targeted at the level of transcription, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, or PCNA. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus

vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[300] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MLH1 or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MLH1. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-MLH1 antibody, e.g., a neutralizing antibody that inactivates MLH1. In still other embodiments, the inhibitor can be a dominant negative mutant of MLH1. In still other embodiments, the inhibitor can be targeted at the level of transcription of MLH1, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MLH1. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[301] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PMS2 (or MutL alpha) or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of PMS2 (or MutL alpha). In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-PMS2 (or MutL alpha) antibody, e.g., a neutralizing antibody that inactivates PMS2 (or MutL alpha). In still other embodiments, the inhibitor can be a dominant negative mutant of PMS2 (or MutL alpha). In still other embodiments, the inhibitor can be targeted at the level of transcription of PMS2 (or MutL alpha), e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding ML PMS2 (or MutL alpha). In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can

include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[302] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PMS1 (or MutL beta) or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of PMS1 (or MutL beta). In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-PMS1 (or MutL beta) antibody, e.g., a neutralizing antibody that inactivates PMS1 (or MutL beta). In still other embodiments, the inhibitor can be a dominant negative mutant of PMS1 (or MutL beta). In still other embodiments, the inhibitor can be targeted at the level of transcription of PMS1 (or MutL beta), e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding PMS1 (or MutL beta). In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[303] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MLH3 (or MutL gamma) or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MLH3 (or MutL gamma). In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-MLH3 (or MutL gamma) antibody, e.g., a neutralizing antibody that

inactivates MLH3 (or MutL gamma). In still other embodiments, the inhibitor can be a dominant negative mutant of MLH3 (or MutL gamma). In still other embodiments, the inhibitor can be targeted at the level of transcription of P MLH3 (or MutL gamma), e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MLH3 (or MutL gamma). In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[304] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MutS alpha (MSH2-MSH6) or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MutS alpha (MSH2-MSH6). In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-MutS alpha (MSH2-MSH6) antibody, e.g., a neutralizing antibody that inactivates MutS alpha (MSH2-MSH6). In still other embodiments, the inhibitor can be a dominant negative mutant of MutS alpha (MSH2-MSH6). In still other embodiments, the inhibitor can be targeted at the level of transcription of MutS alpha (MSH2-MSH6), e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MutS alpha (MSH2-MSH6). In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[305] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise

inactivating MSH2 or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MSH2. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti- MSH2 antibody, e.g., a neutralizing antibody that inactivates MSH2. In still other embodiments, the inhibitor can be a dominant negative mutant of MSH2. In still other embodiments, the inhibitor can be targeted at the level of transcription of MSH2, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MSH2. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[306] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating MSH6 or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of MSH6. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti- MSH6 antibody, e.g., a neutralizing antibody that inactivates MSH6. In still other embodiments, the inhibitor can be a dominant negative mutant of MSH6. In still other embodiments, the inhibitor can be targeted at the level of transcription of MSH6, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding MSH6. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[307] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating PCNA or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of PCNA. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti- PCNA antibody, e.g., a neutralizing antibody that inactivates PCNA. In still other embodiments, the inhibitor can be a dominant negative mutant of PCNA. In still other embodiments, the inhibitor can be targeted at the level of transcription of PCNA, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding PCNA. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[308] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating RFC or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of RFC. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti-RFC antibody, e.g., a neutralizing antibody that inactivates RFC. In still other embodiments, the inhibitor can be a dominant negative mutant of RFC. In still other embodiments, the inhibitor can be targeted at the level of transcription of RFC, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding RFC. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or

lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[309] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating EXO1 or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of EXO1. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti- EXO1 antibody, e.g., a neutralizing antibody that inactivates EXO1. In still other embodiments, the inhibitor can be a dominant negative mutant of EXO1. In still other embodiments, the inhibitor can be targeted at the level of transcription of EXO1, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding EXO1. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[310] In still another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome) using prime editing while blocking, inhibiting, or otherwise inactivating POLδ or variant thereof. In another aspect, the present disclosure provides a method for editing a nucleotide molecule (e.g., a genome), comprising contacting a target nucleotide molecule with a prime editor and an inhibitor of POLδ. In various embodiments, the inhibitor can be a small molecule inhibitor. In other embodiments, the inhibitor can be an anti- POLδ antibody, e.g., a neutralizing antibody that inactivates POLδ. In still other embodiments, the inhibitor can be a dominant negative mutant of POLδ. In still other embodiments, the inhibitor can be targeted at the level of transcription of POLδ, e.g., an siRNA or other nucleic acid agent that knocks down the level of a transcript encoding POLδ. In yet other embodiments, the step of "contacting a target nucleotide molecule with a prime editor" can include (i) delivering directly to a cell an effective amount of a prime editor fusion protein (e.g., PE1 or PE2) complexed with a lipid delivery system; (ii) delivery to a cell a mRNA or delivery complex comprising an mRNA

that encodes a prime editor fusion protein and/or a suitable pegRNA; and (iii) a DNA vector (e.g., an AAV or lentivirus vector, plasmid, or other nucleic acid delivery vector) that encodes a prime editor fusion protein and/or a suitable pegRNA on one or more DNA vectors.

[311] In still other aspects, the present disclosure provides methods for prime editing whereby correction by the MMR pathway of the alterations introduced into a target nucleic acid molecule is evaded, without the need to provide an inhibitor of the MMR pathway. Surprisingly, pegRNAs designed with consecutive nucleotide mismatches compared to a target site on the target nucleic acid, for example, pegRNAs that have three or more consecutive mismatching nucleotides, can evade correction by the MMR pathway, resulting in an increase in prime editing efficiency and/or a decrease in the frequency of indel formation compared to the introduction of a single nucleotide mismatch using prime editing. In addition, insertions and deletions of multiple consecutive nucleotides, for example, three or more contiguous nucleotides, or 10 or more contiguous nucleotides in length introduced by prime editing may also evade correction by the MMR pathway, resulting in an increase in prime editing efficiency and/or a decrease in the frequency of indel formation compared to prime editing with a corresponding control pegRNA (e.g., a control pegRNA that does not introduce insertion or deletion of three or more contiguous nucleotides). In some embodiments, prime editing that introduces insertion or deletion of 10 or more contiguous nucleotides results in an increase in prime editing efficiency and/or a decrease in indel frequency compared to the introduction of an insertion or deletion of less than 10 nucleotides in length using prime editing.

[312] Thus, in one aspect, the present disclosure provides methods for editing a nucleic acid molecule by prime editing comprising contacting a nucleic acid molecule with a prime editor and a pegRNA comprising a DNA synthesis template on its extension arm comprising three or more consecutive nucleotide mismatches relative to a target site on the nucleic acid molecule. In some embodiments, the pegRNA comprises a DNA synthesis template comprising one or more nucleotide edits compared to the endogenous sequence of the nucleic acid molecule (e.g., a double stranded target DNA) to be edited, wherein the one or more nucleotide edits comprises (i) an intended change is an insertion, deletion, or substitution of x consecutive nucleotides that corrects a mutation (e.g. a disease associated mutation) in the nucleic acid molecule, and (ii) an insertion, deletion, or substitution of y consecutive nucleotides directly adjacent to the x nucleotides, wherein (x+y) is an integer no less than 3. In some embodiments, the insertion,

deletion, or substitution of the y consecutive nucleotides is a silent mutation. In some embodiments, the insertion, deletion, or substitution of the y consecutive nucleotides is a benign mutation. The silent mutations may be present in coding regions of the target nucleic acid molecule or in non-coding regions of the target nucleic acid molecule. When the silent mutations are present in a coding region, they introduce into the nucleic acid molecule one or more alternate codons encoding the same amino acid as the unedited nucleic acid molecule. Alternatively, when the silent mutations are in a non-coding region or a junction of a coding region and a non-coding region (e.g., an intron/exon junction), the silent mutations may be present in a region of the nucleic acid molecule that does not influence splicing, gene regulation, RNA lifetime, or other biological properties of the target site on the nucleic acid molecule. A benign mutation may refer to a nucleotide alteration or amino acid alteration that alters the amino acid sequence of the protein or polypeptide encoded by the target nucleic acid sequence, but does not impair or substantively impair expression and/or function of the protein or polypeptide. In some embodiments, x is an integer between 1 and 50. In some embodiments, y is an integer between 1 and 50. In some embodiments, y is an integer no less than 1. In some embodiments, the inclusion of the silent mutation(s) increases the efficiency, reduces unintended indel frequency, and/or improves editing outcome purity by prime editing. As used herein, the term "prime editing outcome purity" may refer to the ratio of intended edit to unintended indels that result from prime editing. In some embodiments, the inclusion of the silent mutation(s) increases the efficiency, reduces unintended indel frequency, and/or improves editing outcome purity by prime editing by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold,

at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold,at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold compared to prime editing with a control pegRNA that does not include the silent mutation(s), e.g., a control pegRNA that only includes the insertion, deletion, or substitution of the x consecutive nucleotides and not the insertion, deletion, or substitution of the y consecutive nucleotides.

[313] In some embodiments, at least one of the three or more consecutive nucleotide mismatches results in an alteration in the amino acid sequence of a protein expressed from the nucleic acid molecule. In some embodiments, more than one of the consecutive nucleotide mismatches results in an alteration in the amino acid sequence of a protein expressed from the nucleic acid molecule. In some embodiments, at least one of the nucleotide mismatches are silent mutations that do not result in an alteration in the amino acid sequence of a protein expressed from the nucleic acid molecule. The silent mutations may be present in coding regions of the target nucleic acid molecule or in non-coding regions of the target nucleic acid molecule. When the silent mutations are present in a coding region, they introduce into the nucleic acid molecule one or more alternate codons encoding the same amino acid as the unedited nucleic acid molecule. Alternatively, when the silent mutations are in a non-coding region, the silent mutations may be present in a region of the nucleic acid molecule that does not influence splicing, gene regulation, RNA lifetime, or other biological properties of the target site on the nucleic acid molecule.

[314] Any number of consecutive nucleotide mismatches of three or more can be used to achieve the benefits of evading correction by the MMR pathway. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 consecutive nucleotide mismatches relative to the endogenous sequence of a target site in the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 3, 4, or 5 consecutive nucleotide mismatches relative to the endogenous sequence of a target site in the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, or 20 consecutive nucleotide mismatches relative to the endogenous sequence

of a target site in the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, or ten or more consecutive nucleotide mismatches relative to a target site on the nucleic acid molecule.

[315] In another aspect, the present disclosure provides methods for editing a nucleic acid molecule by prime editing comprising contacting a nucleic acid molecule with a prime editor and a pegRNA comprising a DNA synthesis template on its extension arm comprising an insertion or deletion of 10 or more nucleotides relative to a target site on the nucleic acid molecule. Insertions and deletions of 10 or more nucleotides in length evade correction by the MMR pathway when introduced by prime editing and thus can benefit from the inhibition of the MMR pathway without the need to provide an inhibitor of MMR. Insertions and deletions of any length greater than 10 nucleotides can be used to achieve the benefits of evading correction by the MMR pathway. In some embodiments, the DNA synthesis template comprises an insertion or deletion of 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides relative to the endogenous sequence at a target site of the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template comprises an insertion or deletion of 11 or more nucleotides, 12 or more nucleotides, 13 or more nucleotides, 14 or more nucleotides, 15 or more nucleotides, 16 or more nucleotides, 17 or more nucleotides, 18 or more nucleotides, 19 or more nucleotides, 20 or more nucleotides, 21 or more nucleotides, 22 or more nucleotides, 23 or more nucleotides, 24 or more nucleotides, or 25 or more nucleotides relative to a target site on a nucleic acid molecule. In certain embodiments, the DNA synthesis template comprises an insertion or deletion of 15 or more nucleotides relative to a target site on the nucleic acid molecule.

[316] The present disclosure provides compositions and methods for prime editing with improved editing efficiency and/or reduced indel formation by inhibiting the DNA mismatch repair pathway while conducting prime editing of a target site. Accordingly, the present disclosure provides a method for editing a nucleic acid molecule by prime editing that involves contacting a nucleic acid molecule with a prime editor, a pegRNA, and an inhibitor of the DNA mismatch repair pathway, thereby installing one or more modifications to the nucleic acid molecule at a target site with increased editing efficiency and/or lower indel formation. The present disclosure further provides polynucleotides for editing a DNA target site by prime

editing comprising a nucleic acid sequence encoding a napDNAbp, a polymerase, and an inhibitor of the DNA mismatch repair pathway, wherein the napDNAbp and polymerase is capable in the presence of a pegRNA of installing one or more modifications in the DNA target site with increased editing efficiency and/or lower indel formation. Thus, the methods and compositions described herein utilize prime editors, which may comprise a nucleic acid programmable DNA binding protein (napDNAbp).

Prime editors: napDNAbp domain

[317] In one aspect, a napDNAbp of the prime editors described herein can be associated with or complexed with at least one guide nucleic acid (e.g., guide RNA or a PEgRNA), which localizes the napDNAbp to a DNA sequence that comprises a DNA strand (i.e., a target strand) that is complementary to the guide nucleic acid, or a portion thereof (e.g., the spacer of a guide RNA which anneals to the protospacer of the DNA target). In other words, the guide nucleic-acid "programs" the napDNAbp (e.g., Cas9 or equivalent) to localize and bind to complementary sequence of the protospacer in the DNA.

[318] Any suitable napDNAbp may be used in the prime editors utilized in the methods and compositions described herein. In various embodiments, the napDNAbp may be any Class 2 CRISPR-Cas system, including any type II, type V, or type VI CRISPR-Cas enzyme. Given the rapid development of CRISPR-Cas as a tool for genome editing, there have been constant developments in the nomenclature used to describe and/or identify CRISPR-Cas enzymes, such as Cas9 and Cas9 orthologs. This application references CRISPR-Cas enzymes with nomenclature that may be old and/or new. The skilled person will be able to identify the specific CRISPR-Cas enzyme being referenced in this Application based on the nomenclature that is used, whether it is old (i.e., "legacy") or new nomenclature. CRISPR-Cas nomenclature is extensively discussed in Makarova *et al.*, "Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?," *The CRISPR Journal*, Vol. 1. No. 5, 2018, the entire contents of which are incorporated herein by reference. The particular CRISPR-Cas nomenclature used in any given instance in this Application is not limiting in any way and the skilled person will be able to identify which CRISPR-Cas enzyme is being referenced.

[319] For example, the following type II, type V, and type VI Class 2 CRISPR-Cas enzymes have the following art-recognized old (i.e., legacy) and new names. Each of these enzymes,

and/or variants thereof, may be used with the prime editors utilized in the methods and compositions described herein:

Legacy nomenclature	Current nomenclature*
type II CRISPR-Cas enzymes	
Cas9	same
type V CRISPR-Cas enzymes	
Cpf1	Cas12a
CasX	Cas12e
C2c1	Cas12b1
Cas12b2	same
C2c3	Cas12c
CasY	Cas12d
C2c4	same
C2c8	same
C2c5	same
C2c10	same
C2c9	same
type VI CRISPR-Cas enzymes	
C2c2	Cas13a
Cas13d	same
C2c7	Cas13c
C2c6	Cas13b

* See Makarova *et al.*, *The CRISPR Journal*, Vol. 1, No. 5, 2018

[320] Without being bound by theory, the mechanism of action of certain napDNAbp contemplated herein includes the step of forming an R-loop whereby the napDNAbp induces the unwinding of a double-strand DNA target, thereby separating the strands in the region bound by the napDNAbp. The guide RNA spacer then hybridizes to the "target strand" at a region that is complementary to the protospacer sequence. This displaces a "non-target strand" that is complementary to the target strand, which forms the single strand region of the R-loop. In some embodiments, the napDNAbp includes one or more nuclease activities, which then cut the DNA leaving various types of lesions. For example, the napDNAbp may comprises a nuclease activity that cuts the non-target strand at a first location, and/ or cuts the target strand at a second location. Depending on the nuclease activity, the target DNA can be cut to form a "double-stranded break" whereby both strands are cut. In other embodiments, the target DNA can be cut at only a single site, i.e., the DNA is "nicked" on one strand. Exemplary napDNAbp with different nuclease activities include "Cas9 nickase" ("nCas9") and a deactivated Cas9 having no nuclease activities ("dead Cas9" or "dCas9").

[321] The below description of various napDNAbps which can be used in connection with the prime editors utilized in the presently disclosed methods and compositions is not meant to be

limiting in any way. The prime editors may comprise the canonical SpCas9, or any ortholog Cas9 protein, or any variant Cas9 protein —including any naturally occurring variant, mutant, or otherwise engineered version of Cas9 — that is known or that can be made or evolved through a directed evolutionary or otherwise mutagenic process. In various embodiments, the Cas9 or Cas9 variants have a nickase activity, i.e., only cleave one strand of the target DNA sequence. In other embodiments, the Cas9 or Cas9 variants have inactive nucleases, i.e., are "dead" Cas9 proteins. Other variant Cas9 proteins that may be used are those having a smaller molecular weight than the canonical SpCas9 (e.g., for easier delivery) or having modified or rearranged primary amino acid structure (e.g., the circular permutant formats).

[322] The prime editors utilized in the methods and compositions described herein may also comprise Cas9 equivalents, including Cas12a (Cpf1) and Cas12b1 proteins which are the result of convergent evolution. The napDNAbps used herein (e.g., SpCas9, Cas9 variant, or Cas9 equivalents) may also contain various modifications that alter/enhance their PAM specificities. Lastly, the application contemplates any Cas9, Cas9 variant, or Cas9 equivalent which has at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.9% sequence identity to a reference Cas9 sequence, such as a reference SpCas9 canonical sequence or a reference Cas9 equivalent (e.g., Cas12a (Cpf1)).

[323] The napDNAbp can be a CRISPR (clustered regularly interspaced short palindromic repeat)-associated nuclease. As outlined above, CRISPR is an adaptive immune system that provides protection against mobile genetic elements (viruses, transposable elements, and conjugative plasmids). CRISPR clusters contain spacers, sequences complementary to antecedent mobile elements, and target invading nucleic acids. CRISPR clusters are transcribed and processed into CRISPR RNA (crRNA). In type II CRISPR systems, correct processing of pre-crRNA requires a trans-encoded small RNA (tracrRNA), endogenous ribonuclease 3 (rnc), and a Cas9 protein. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves a linear or circular dsDNA target complementary to the spacer. The target strand not complementary to crRNA is first cut endonucleolytically, then trimmed 3′-5′ exonucleolytically. In nature, DNA-binding and cleavage typically requires protein and both RNAs. However, single guide RNAs ("sgRNA", or simply "gRNA") can be engineered so as to incorporate aspects of both the crRNA and

tracrRNA into a single RNA species. See, e.g., Jinek M. *et al.*, *Science* 337:816-821(2012), the entire contents of which is hereby incorporated by reference.

[324] In some embodiments, the napDNAbp directs cleavage of one or both strands at the location of a target sequence, such as within the target sequence and/or within the complement of the target sequence. In some embodiments, the napDNAbp directs cleavage of one or both strands within about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 50, 100, 200, 500, or more base pairs from the first or last nucleotide of a target sequence. In some embodiments, a vector encodes a napDNAbp that is mutated to with respect to a corresponding wild-type enzyme such that the mutated napDNAbp lacks the ability to cleave one or both strands of a target polynucleotide containing a target sequence. For example, an aspartate-to-alanine substitution (D10A) in the RuvC I catalytic domain of Cas9 from *S. pyogenes* converts Cas9 from a nuclease that cleaves both strands to a nickase (cleaves a single strand). Other examples of mutations that render Cas9 a nickase include, without limitation, H840A, N854A, and N863A in reference to the canonical SpCas9 sequence, or to equivalent amino acid positions in other Cas9 variants or Cas9 equivalents.

[325] As used herein, the term "Cas protein" refers to a full-length Cas protein obtained from nature, a recombinant Cas protein having a sequences that differs from a naturally occurring Cas protein, or any fragment of a Cas protein that nevertheless retains all or a significant amount of the requisite basic functions needed for the disclosed methods, i.e., (i) possession of nucleic-acid programmable binding of the Cas protein to a target DNA, and (ii) ability to nick the target DNA sequence on one strand. The Cas proteins contemplated herein embrace CRISPR Cas 9 proteins, as well as Cas9 equivalents, variants (e.g., Cas9 nickase (nCas9) or nuclease inactive Cas9 (dCas9)) homologs, orthologs, or paralogs, whether naturally occurring or non-naturally occurring (e.g., engineered or recombinant), and may include a Cas9 equivalent from any Class 2 CRISPR system (e.g., type II, V, VI), including Cas12a (Cpf1), Cas12e (CasX), Cas12b1 (C2c1), Cas12b2, Cas12c (C2c3), C2c4, C2c8, C2c5, C2c10, C2c9 Cas13a (C2c2), Cas13d, Cas13c (C2c7), Cas13b (C2c6), and Cas13b. Further Cas-equivalents are described in Makarova *et al.*, "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector," Science 2016; 353(6299) and Makarova *et al.*, "Classification and Nomenclature of CRISPR-Cas Systems: Where from Here?," *The CRISPR Journal*, Vol. 1. No. 5, 2018, the contents of which are incorporated herein by reference.

[326] The terms "Cas9" or "Cas9 nuclease" or "Cas9 moiety" or "Cas9 domain" embrace any naturally occurring Cas9 from any organism, any naturally-occurring Cas9 equivalent or functional fragment thereof, any Cas9 homolog, ortholog, or paralog from any organism, and any mutant or variant of a Cas9, naturally-occurring or engineered. The term Cas9 is not meant to be particularly limiting and may be referred to as a "Cas9 or equivalent." Exemplary Cas9 proteins are further described herein and/or are described in the art and are incorporated herein by reference. The present disclosure is unlimited with regard to the particular Cas9 that is employed in the prime editors utilized in the methods and compositions described herein.

[327] As noted herein, Cas9 nuclease sequences and structures are well-known to those of skill in the art (see, e.g., "Complete genome sequence of an M1 strain of Streptococcus pyogenes." Ferretti *et al.*, J.J., McShan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K., Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P., Qian Y., Jia H.G., Najar F.Z., Ren Q., Zhu H., Song L., White J., Yuan X., Clifton S.W., Roe B.A., McLaughlin R.E., *Proc. Natl. Acad. Sci.* U.S.A. 98:4658-4663(2001); "CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III." Deltcheva E., Chylinski K., Sharma C.M., Gonzales K., Chao Y., Pirzada Z.A., Eckert M.R., Vogel J., Charpentier E., Nature 471:602-607(2011); and "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity." Jinek M., Chylinski K., Fonfara I., Hauer M., Doudna J.A., Charpentier E. *Science* 337:816-821(2012), the entire contents of each of which are incorporated herein by reference).

[328] Examples of Cas9 and Cas9 equivalents are provided as follows; however, these specific examples are not meant to be limiting. The prime editors utilized in the methods and compositions of the present disclosure may use any suitable napDNAbp, including any suitable Cas9 or Cas9 equivalent.

A. Wild type canonical SpCas9

[329] In one embodiment, the prime editor constructs utilized in the methods and compositions described herein may comprise the "canonical SpCas9" nuclease from *S. pyogenes*, which has been widely used as a tool for genome engineering and is categorized as the type II subgroup of enzymes of the Class 2 CRISPR-Cas systems. This Cas9 protein is a large, multi-domain protein containing two distinct nuclease domains. Point mutations can be introduced into Cas9 to abolish one or both nuclease activities, resulting in a nickase Cas9 (nCas9) or dead Cas9 (dCas9), respectively, that still retains its ability to bind DNA in a sgRNA-programmed manner. In

principle, when fused to another protein or domain, Cas9, or variant thereof (e.g., nCas9) can target that protein to virtually any DNA sequence simply by co-expression with an appropriate sgRNA. As used herein, the canonical SpCas9 protein refers to the wild type protein from *Streptococcus pyogenes* having the following amino acid sequence:

Description	Sequence	SEQ ID NO:
SpCas9 *Streptococcus pyogenes* M1 SwissProt Accession No. Q99ZW2 Wild type	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD SGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEE DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIK FRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDD DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDN REKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSF IERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQ KKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKR RRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDI QKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYL QNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDN VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLV ETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVR EINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEI GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR KVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPT VAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKK DLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKP IREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYET RIDLSQLGGD	2
SpCas9 Reverse translation of SwissProt Accession No. Q99ZW2 *Streptococcus pyogenes*	ATGGATAAAAAATATAGCATTGGCCTGGATATTGGCACCAACAGCGTGGGC TGGGCGGTGATTACCGATGAATATAAAGTGCCGAGCAAAAAATTTAAAGTG CTGGGCAACACCGATCGCCATAGCATTAAAAAAAACCTGATTGGCGCGCTG CTGTTTGATAGCGGCGAAACCGCGGAAGCGACCCGCCTGAAACGCACCGCG CGCCGCCGCTATACCCGCCGCAAAAACCGCATTTGCTATCTGCAGGAAATTT TTAGCAACGAAATGGCGAAAGTGGATGATAGCTTTTTTCATCGCCTGGAAGA AAGCTTTCTGGTGGAAGAAGATAAAAAACATGAACGCCATCCGATTTTTGG CAACATTGTGGATGAAGTGGCGTATCATGAAAAATATCCGACCATTTATCAT CTGCGCAAAAAACTGGTGGATAGCACCGATAAAGCGGATCTGCGCCTGATT TATCTGGCGCTGGCGCATATGATTAAATTTCGCGGCCATTTTCTGATTGAAG GCGATCTGAACCCGGATAACAGCGATGTGGATAAACTGTTTATTCAGCTGGT GCAGACCTATAACCAGCTGTTTGAAGAAAACCCGATTAACGCGAGCGGCGT GGATGCGAAAGCGATTCTGAGCGCGCGCCTGAGCAAAAGCCGCCGCCTGGA AAACCTGATTGCGCAGCTGCCGGGCGAAAAAAAAAAACGGCCTGTTTGGCAA CCTGATTGCGCTGAGCCTGGGCCTGACCCCGAACTTTAAAAGCAACTTTGAT CTGGCGGAAGATGCGAAACTGCAGCTGAGCAAAGATACCTATGATGATGAT CTGGATAACCTGCTGGCGCAGATTGGCGATCAGTATGCGGATCTGTTTCTGG CGGCGAAAAACCTGAGCGATGCGATTCTGCTGAGCGATATTCTGCGCGTGA ACACCGAAATTACCAAAGCGCCGCTGAGCGCGAGCATGATTAAACGCTATG ATGAACATCATCAGGATCTGACCCTGCTGAAAGCGCTGGTGCGCCAGCAGC	3

	TGCCGGAAAAATATAAAGAAATTTTTTTTGATCAGAGCAAAAACGGCTATG CGGGCTATATTGATGGCGGCGCGAGCCAGGAAGAATTTTATAAATTTATTAA ACCGATTCTGGAAAAAATGGATGGCACCGAAGAACTGCTGGTGAAACTGAA CCGCGAAGATCTGCTGCGCAAACAGCGCACCTTTGATAACGGCAGCATTCC GCATCAGATTCATCTGGGCGAACTGCATGCGATTCTGCGCCGCCAGGAAGAT TTTTATCCGTTTCTGAAAGATAACCGCGAAAAAATTGAAAAAATTCTGACCT TTCGCATTCCGTATTATGTGGGCCCGCTGGCGCGCGGCAACAGCCGCTTTGC GTGGATGACCCGCAAAAGCGAAGAAACCATTACCCCGTGGAACTTTGAAGA AGTGGTGGATAAAGGCGCGAGCGCGCAGAGCTTTATTGAACGCATGACCAA CTTTGATAAAAACCTGCCGAACGAAAAAGTGCTGCCGAAACATAGCCTGCT GTATGAATATTTTACCGTGTATAACGAACTGACCAAAGTGAAATATGTGACC GAAGGCATGCGCAAACCGGCGTTTCTGAGCGGCGAACAGAAAAAAGCGATT GTGGATCTGCTGTTTAAAACCAACCGCAAAGTGACCGTGAAACAGCTGAAA GAAGATTATTTTAAAAAAATTGAATGCTTTGATAGCGTGGAAATTAGCGGCG TGGAAGATCGCTTTAACGCGAGCCTGGGCACCTATCATGATCTGCTGAAAAT TATTAAAGATAAAGATTTTCTGGATAACGAAGAAAACGAAGATATTCTGGA AGATATTGTGCTGACCCTGACCCTGTTTGAAGATCGCGAAATGATTGAAGAA CGCCTGAAAACCTATGCGCATCTGTTTGATGATAAAGTGATGAAACAGCTGA AACGCCGCCGCTATACCGGCTGGGGCCGCCTGAGCCGCAAACTGATTAACG GCATTCGCGATAAACAGAGCGGCAAAACCATTCTGGATTTTCTGAAAAGCG ATGGCTTTGCGAACCGCAACTTTATGCAGCTGATTCATGATGATAGCCTGAC CTTTAAAGAAGATATTCAGAAAGCGCAGGTGAGCGGCCAGGGCGATAGCCT GCATGAACATATTGCGAACCTGGCGGGCAGCCCGGCGATTAAAAAAGGCAT TCTGCAGACCGTGAAAGTGGTGGATGAACTGGTGAAAGTGATGGGCCGCCA TAAACCGGAAAACATTGTGATTGAAATGGCGCGCGAAAACCAGACCACCCA GAAAGGCCAGAAAAACAGCCGCGAACGCATGAAACGCATTGAAGAAGGCA TTAAAGAACTGGGCAGCCAGATTCTGAAAGAACATCCGGTGGAAAACACCC AGCTGCAGAACGAAAAACTGTATCTGTATTATCTGCAGAACGGCCGCGATA TGTATGTGGATCAGGAACTGGATATTAACCGCCTGAGCGATTATGATGTGGA TCATATTGTGCCGCAGAGCTTTCTGAAAGATGATAGCATTGATAACAAAGTG CTGACCCGCAGCGATAAAAACCGCGGCAAAAGCGATAACGTGCCGAGCGAA GAAGTGGTGAAAAAAATGAAAAACTATTGGCGCCAGCTGCTGAACGCGAAA CTGATTACCCAGCGCAAATTTGATAACCTGACCAAAGCGGAACGCGGCGGC CTGAGCGAACTGGATAAAGCGGGCTTTATTAAACGCCAGCTGGTGGAAACC CGCCAGATTACCAAACATGTGGCGCAGATTCTGGATAGCCGCATGAACACC AAATATGATGAAAACGATAAACTGATTCGCGAAGTGAAAGTGATTACCCTG AAAAGCAAACTGGTGAGCGATTTTCGCAAAGATTTTCAGTTTTATAAAGTGC GCGAAATTAACAACTATCATCATGCGCATGATGCGTATCTGAACGCGGTGGT GGGCACCGCGCTGATTAAAAAATATCCGAAACTGGAAAGCGAATTTGTGTA TGGCGATTATAAAGTGTATGATGTGCGCAAAATGATTGCGAAAAGCGAACA GGAAATTGGCAAAGCGACCGCGAAATATTTTTTTTATAGCAACATTATGAAC TTTTTTAAAACCGAAATTACCCTGGCGAACGGCGAAATTCGCAAACGCCCGC TGATTGAAACCAACGGCGAAACCGGCGAAATTGTGTGGGATAAAGGCCGCG ATTTTGCGACCGTGCGCAAAGTGCTGAGCATGCCGCAGGTGAACATTGTGA AAAAAACCGAAGTGCAGACCGGCGGCTTTAGCAAAGAAAGCATTCTGCCGA AACGCAACAGCGATAAACTGATTGCGCGCAAAAAAGATTGGGATCCGAAAA AATATGGCGGCTTTGATAGCCCGACCGTGGCGTATAGCGTGCTGGTGGTGGC GAAAGTGGAAAAAGGCAAAAGCAAAAAACTGAAAAGCGTGAAAGAACTGC TGGGCATTACCATTATGGAACGCAGCAGCTTTGAAAAAAACCCGATTGATTT TCTGGAAGCGAAAGGCTATAAAGAAGTGAAAAAAGATCTGATTATTAAACT GCCGAAATATAGCCTGTTTGAACTGGAAAACGGCCGCAAACGCATGCTGGC GAGCGCGGGCGAACTGCAGAAAGGCAACGAACTGGCGCTGCCGAGCAAAT ATGTGAACTTTCTGTATCTGGCGAGCCATTATGAAAAACTGAAAGGCAGCCC GGAAGATAACGAACAGAAACAGCTGTTTGTGGAACAGCATAAACATTATCT GGATGAAATTATTGAACAGATTAGCGAATTTAGCAAACGCGTGATTCTGGC	

	GGATGCGAACCTGGATAAAGTGCTGAGCGCGTATAACAAACATCGCGATAA ACCGATTCGCGAACAGGCGGAAAACATTATTCATCTGTTTACCCTGACCAAC CTGGGCGCGCCGGCGGCGTTTAAATATTTTGATACCACCATTGATCGCAAAC GCTATACCAGCACCAAAGAAGTGCTGGATGCGACCCTGATTCATCAGAGCA TTACCGGCCTGTATGAAACCCGCATTGATCTGAGCCAGCTGGGCGGCGAT	

[330] The prime editors utilized in the methods and compositions described herein may include canonical SpCas9, or any variant thereof having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity with a wild type Cas9 sequence provided above. These variants may include SpCas9 variants containing one or more mutations, including any known mutation reported with the SwissProt Accession No. Q99ZW2 (SEQ ID NO: 2) entry, which include:

SpCas9 mutation (relative to the amino acid sequence of the canonical SpCas9 sequence, SEQ ID NO: 2)	Function/Characteristic (as reported) (see UniProtKB – Q99ZW2 (CAS9_STRPT1) entry – incorporated herein by reference)
D10A	Nickase mutant which cleaves the protospacer strand (but no cleavage of non-protospacer strand)
S15A	Decreased DNA cleavage activity
R66A	Decreased DNA cleavage activity
R70A	No DNA cleavage
R74A	Decreased DNA cleavage
R78A	Decreased DNA cleavage
97-150 deletion	No nuclease activity
R165A	Decreased DNA cleavage
175-307 deletion	About 50% decreased DNA cleavage
312-409 deletion	No nuclease activity
E762A	Nickase
H840A	Nickase mutant which cleaves the non-protospacer strand but does not cleave the protospacer strand
N854A	Nickase
N863A	Nickase
H982A	Decreased DNA cleavage
D986A	Nickase
1099-1368 deletion	No nuclease activity
R1333A	Reduced DNA binding

[331] Other wild type SpCas9 sequences that may be used in the present disclosure, include:

Description	Sequence	SEQ ID NO:
SpCas9 *Streptococcus pyogenes* MGAS1882 wild type NC_017053.1	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGAT GGGCGGTGATCACTGATGATTATAAGGTTCCGTCTAAAAAGTTCAAGGTTCT GGGAAATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTTTA TTTGGCAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTA GAAGGTATACACGTCGGAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCA AATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGTCTTT TTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATA GTAGATGAAGTTGCTTATCATGAGAAATATCCAACTATCTATCATCTGCGAAA AAAATTGGCAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT TAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAAT CCTGATAATAGTGATGTGGACAAACTATTTATCCAGTTGGTACAAATCTACAA	4

	TCAATTATTTGAAGAAAACCCTATTAACGCAAGTAGAGTAGATGCTAAAGCG ATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTCA GCTCCCCGGTGAGAAGAGAAATGGCTTGTTTGGGAATCTCATTGCTTTGTCAT TGGGATTGACCCCTAATTTTAAATCAAATTTTGATTTGGCAGAAGATGCTAAA TTACAGCTTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGCA AATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGATG CTATTTTACTTTCAGATATCCTAAGAGTAAATAGTGAAATAACTAAGGCTCCC CTATCAGCTTCAATGATTAAGCGCTACGATGAACATCATCAAGACTTGACTCT TTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAAATCTTTT TTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCCA AGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGATGGTACT GAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGGA CCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGTGAGCTGCATGCT ATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAATCGTGAGA AGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCG CGTGGCAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTA CCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTT ATTGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTAC CAAAACATAGTTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAG GTCAAATATGTTACTGAGGGAATGCGAAAACCAGCATTTCTTTCAGGTGAAC AGAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGT TAAGCAATTAAAAGAAGATTATTTCAAAAAAATAGAATGTTTTGATAGTGTT GAAATTTCAGGAGTTGAAGATAGATTTAATGCTTCATTAGGCGCCTACCATG ATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGA AGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGGGA TGATTGAGGAAAGACTTAAAACATATGCTCACCTCTTTGATGATAAGGTGAT GAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAA TTGATTAATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTT GAAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATA GTTTGACATTTAAAGAAGATATTCAAAAAGCACAGGTGTCTGGACAAGGCCA TAGTTTACATGAACAGATTGCTAACTTAGCTGGCAGTCCTGCTATTAAAAAAG GTATTTTACAGACTGTAAAAATTGTTGATGAACTGGTCAAAGTAATGGGGCA TAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGACAACTCAA AAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGTATC AAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTCAAT TGCAAAATGAAAAGCTCTATCTCTATTATCTACAAAATGGAAGAGACATGTA TGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGATCACA TTGTTCCACAAAGTTTCATTAAAGACGATTCAATAGACAATAAGGTACTAAC GCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTA GTCAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAATCA CTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGA ACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCA CTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGA AAATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTA GTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAA TTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGA TTAAGAAATATCCAAAACTTGAATCGGAGTTTGTCTATGGTGATTATAAAGTT TATGATGTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGCAAAGCAA CCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAAATT ACACTTGCAAATGGAGAGATTCGCAAACGCCCTCTAATCGAAACTAATGGGG AAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACAGTGCGCAA AGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAGACA GGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGCTTA TTGCTCGTAAAAAAGACTGGGATCCAAAAAAATATGGTGGTTTTGATAGTCC AACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAATCG	

	AAGAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAAGAA GTTCCTTTGAAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAAGGA AGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAGTTAG AAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGGAA ATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGTCAT TATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTG TGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATT TTCTAAGCGTGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGCAT ATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATATTATTC ATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATA CAACAATTGATCGTAAACGATATACGTCTACAAAAGAAGTTTTAGATGCCAC TCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGCATTGATTTGAGTC AGCTAGGAGGTGACTGA	
SpCas9 *Streptococcus pyogenes* MGAS1882 wild type NC_017053.1	MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLADSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQIYNQLFEENPINASRVDAKAILSARLSKSRR LENLIAQLPGEKRNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNSEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGAYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDRGMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG HSLHEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTQKG QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQE LDINRLSDYDVDHIVPQSFIKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNY WRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDS RMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNA VVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNF FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEV QTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRK RMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKH YLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAP AAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	5
SpCas9 *Streptococcus pyogenes* wild type SWBC2D7W014	ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCGTTGGATG GGCTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTG GGGAACACAGACCGTCATTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATT CGATAGTGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAG AAGGTATACACGTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTTAGC AATGAGATGGCCAAAGTTGACGATTCTTTCTTTCACCGTTTGGAAGAGTCCTT CCTTGTCGAAGAGGACAAGAAACATGAACGGCACCCCATCTTTGGAAACATA GTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTATCACCTCAGAA AAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGC TCTTGCCCATATGATAAAGTTCCGTGGGCACTTTCTCATTGAGGGTGATCTAA ATCCGGACAACTCGGATGTCGACAAACTGTTCATCCAGTTAGTACAAACCTA TAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAG GCTATTCTTAGCGCCCGCCTCTCTAAATCCCGACGGCTAGAAAACCTGATCGC ACAATTACCCGGAGAGAAGAAAAATGGGTTGTTCGGTAACCTTATAGCGCTC TCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAAGATGC CAAATTGCAGCTTAGTAAGGACACGTACGATGACGATCTCGACAATCTACTG GCACAAATTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAAACCTTA	6

	GCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAA GGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGATGAACATCACCAAGAC TTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAGAAATATAAGG AAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGG AGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGATG GATGGGACGGAAGAGTTGCTTGTAAAACTCAATCGCGAAGATCTACTGCGAA AGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAAATCCACTTAGGCGA ATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCCTCAAAGACA ATCGTGAAAAGATTGAGAAAATCCTAACCTTTCGCATACCTTACTATGTGGG ACCCCTGGCCCGAGGGAACTCTCGGTTCGCATGGATGACAAGAAAGTCCGAA GAAACGATTACTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAG CTCAATCGTTCATCGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGA AAAAGTATTGCCTAAGCACAGTTTACTTTACGAGTATTTCACAGTGTACAATG AACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCTTTCT AAGCGGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGACCAACCG CAAAGTGACAGTTAAGCAATTGAAAGAGGACTACTTTAAGAAAATTGAATGC TTCGATTCTGTCGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCACTTGG TACGTATCATGACCTCCTAAAGATAATTAAAGATAAGGACTTCCTGGATAAC GAAGAGAATGAAGATATCTTAGAAGATATAGTGTTGACTCTTACCCTCTTTGA AGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTCGAC GATAAGGTTATGAAACAGTTAAAGAGGCGTCGCTATACGGGCTGGGGACGAT TGTCGCGGAAACTTATCAACGGGATAAGAGACAAGCAAAGTGGTAAAACTAT TCTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGCAGCTGA TCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAGGCACAGGTTTC CGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTCGCCA GCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTA AGGTCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCACGCG AAAATCAAACGACTCAGAAGGGGCAAAAAAACAGTCGAGAGCGGATGAAGA GAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATC CTGTGGAAAATACCCAATTGCAGAACGAGAAACTTTACCTCTATTACCTACA AAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAACCGTTTATCT GATTACGACGTCGATCACATTGTACCCCAATCCTTTTTGAAGGACGATTCAAT CGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAA TGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTC CTAAATGCGAAACTGATAACGCAAAGAAAGTTCGATAACTTAACTAAAGCTG AGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAAACGTCAGCT CGTGGAAACCCGCCAAATCACAAAGCATGTTGCACAGATACTAGATTCCCGA ATGAATACGAAATACGACGAGAACGATAAGCTGATTCGGGAAGTCAAAGTA ATCACTTTAAAGTCAAAATTGGTGTCGGACTTCAGAAAGGATTTTCAATTCTA TAAAGTTAGGGAGATAAATAACTACCACCATGCGCACGACGCTTATCTTAAT GCCGTCGTAGGGACCGCACTCATTAAGAAATACCCGAAGCTAGAAAGTGAGT TTGTGTATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAG CGAACAGGAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATT ATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACGGAGAGATACGCAAAC GACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGG CCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATA GTAAAGAAAACTGAGGTGCAGACCGGAGGGTTTTCAAAGGAATCGATTCTTC CAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGACCCGA AAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCCTAGTAGTG GCAAAAGTTGAGAAGGGAAAATCCAAGAAACTGAAGTCAGTCAAAGAATTA TTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACCCCATCGACTT CCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATTAAACTA CCAAAGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTA GCGCCGGAGAGCTTCAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGT GAATTTCCTGTATTTAGCGTCCCATTACGAGAAGTTGAAAGGTTCACCTGAAG	

	ATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGA AATCATAGAGCAAATTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCC AATCTGGACAAAGTATTAAGCGCATACAACAAGCACAGGGATAAACCCATAC GTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCGCT CCAGCCGCATTCAAGTATTTTGACACAACGATAGATCGCAAACGATACACTT CTACCAAGGAGGTGCTAGACGCGACACTGATTCACCAATCCATCACGGGATT ATATGAAACTCGGATAGATTTGTCACAGCTTGGGGGTGACGGATCCCCCAAG AAGAAGAGGAAAGTCTCGAGCGACTACAAAGACCATGACGGTGATTATAAA GATCATGACATCGATTACAAGGATGACGATGACAAGGCTGCAGGA	
SpCas9 *Streptococcus pyogenes* wild type Encoded product of SWBC2D7W014	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGSPKKKRK VSSDYKDHDGDYKDHDIDYKDDDDKAAG	7
SpCas9 *Streptococcus pyogenes* M1GAS wild type NC_002737.2	ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGAT GGGCGGTGATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCT GGGAAATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTTTA TTTGACAGTGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTA GAAGGTATACACGTCGGAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCA AATGAGATGGCGAAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGTCTTT TTTGGTGGAAGAAGACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATA GTAGATGAAGTTGCTTATCATGAGAAATATCCAACTATCTATCATCTGCGAAA AAAATTGGTAGATTCTACTGATAAAGCGGATTTGCGCTTAATCTATTTGGCCT TAGCGCATATGATTAAGTTTCGTGGTCATTTTTTGATTGAGGGAGATTTAAAT CCTGATAATAGTGATGTGGACAAACTATTTATCCAGTTGGTACAAACCTACA ATCAATTATTTGAAGAAAACCCTATTAACGCAAGTGGAGTAGATGCTAAAGC GATTCTTTCTGCACGATTGAGTAAATCAAGACGATTAGAAAATCTCATTGCTC AGCTCCCCGGTGAGAAGAAAAATGGCTTATTTGGGAATCTCATTGCTTTGTCA TTGGGTTTGACCCCTAATTTTAAATCAAATTTTGATTTGGCAGAAGATGCTAA ATTACAGCTTTCAAAAGATACTTACGATGATGATTTAGATAATTTATTGGCGC AAATTGGAGATCAATATGCTGATTTGTTTTTGGCAGCTAAGAATTTATCAGAT GCTATTTTACTTTCAGATATCCTAAGAGTAAATACTGAAATAACTAAGGCTCC CCTATCAGCTTCAATGATTAAACGCTACGATGAACATCATCAAGACTTGACTC TTTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATAAAGAAATCTTT	8

	TTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGGAGCTAGCC AAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGATGGTAC TGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAACGG ACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGTGAGCTGCATGC TATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAATCGTGAG AAGATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGC GCGTGGCAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATT ACCCCATGGAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATT TATTGAACGCATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTA CCAAAACATAGTTTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAA GGTCAAATATGTTACTGAAGGAATGCGAAAACCAGCATTTCTTTCAGGTGAA CAGAAGAAAGCCATTGTTGATTTACTCTTCAAAACAAATCGAAAAGTAACCG TTAAGCAATTAAAAGAAGATTATTTCAAAAAAATAGAATGTTTTGATAGTGT TGAAATTTCAGGAGTTGAAGATAGATTTAATGCTTCATTAGGTACCTACCATG ATTTGCTAAAAATTATTAAAGATAAAGATTTTTTGGATAATGAAGAAAATGA AGATATCTTAGAGGATATTGTTTTAACATTGACCTTATTTGAAGATAGGGAGA TGATTGAGGAAAGACTTAAAACATATGCTCACCTCTTTGATGATAAGGTGAT GAAACAGCTTAAACGTCGCCGTTATACTGGTTGGGGACGTTTGTCTCGAAAA TTGATTAATGGTATTAGGGATAAGCAATCTGGCAAAACAATATTAGATTTTTT GAAATCAGATGGTTTTGCCAATCGCAATTTTATGCAGCTGATCCATGATGATA GTTTGACATTTAAAGAAGACATTCAAAAAGCACAAGTGTCTGGACAAGGCGA TAGTTTACATGAACATATTGCAAATTTAGCTGGTAGCCCTGCTATTAAAAAAG GTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGTCAAAGTAATGGGGCG GCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAGACAACT CAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGAAGGT ATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATACTC AATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGACAT GTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGATC ACATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAGACAATAAGGTCTTA ACGCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAG TAGTCAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAAT CACTCAACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGT GAACTTGATAAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAA TCACTAAGCATGTGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGA TGAAAATGATAAACTTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAA TTAGTTTCTGACTTCCGAAAAGATTTCCAATTCTATAAAGTACGTGAGATTAA CAATTACCATCATGCCCATGATGCGTATCTAAATGCCGTCGTTGGAACTGCTT TGATTAAGAAATATCCAAAACTTGAATCGGAGTTTGTCTATGGTGATTATAAA GTTTATGATGTTCGTAAAATGATTGCTAAGTCTGAGCAAGAAATAGGCAAAG CAACCGCAAAATATTTCTTTTACTCTAATATCATGAACTTCTTCAAAACAGAA ATTACACTTGCAAATGGAGAGATTCGCAAACGCCCTCTAATCGAAACTAATG GGGAAACTGGAGAAATTGTCTGGGATAAAGGGCGAGATTTTGCCACAGTGCG CAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAAGAAAACAGAAGTACAG ACAGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAGAAATTCGGACAAGC TTATTGCTCGTAAAAAAGACTGGGATCCAAAAAAATATGGTGGTTTTGATAG TCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAAAAAGGGAAA TCGAAGAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATTATGGAAA GAAGTTCCTTTGAAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGATATAA GGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGAGT TAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAG GAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGT CATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGT TTGTGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGA ATTTTCTAAGCGTGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTG CATATAACAAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATATTA	

	TTCATTTATTTACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTG ATACAACAATTGATCGTAAACGATATACGTCTACAAAAGAAGTTTTAGATGC CACTCTTATCCATCAATCCATCACTGGTCTTTATGAAACACGCATTGATTTGA GTCAGCTAGGAGGTGACTGA	
SpCas9 *Streptococcus pyogenes* M1GAS wild type Encoded product of NC_002737.2 (100% identical to the canonical Q99ZW2 wild type)	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	2

[332] The prime editors utilized in the methods and compositions described herein may include any of the above SpCas9 sequences, or any variant thereof having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto.

B. Wild type Cas9 orthologs

[333] In other embodiments, the Cas9 protein can be a wild type Cas9 ortholog from another bacterial species different from the canonical Cas9 from *S. pyogenes*. For example, the following Cas9 orthologs can be used in connection with the prime editor constructs utilized in the methods and compositions described in this specification. In addition, any variant Cas9 orthologs having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity to any of the below orthologs may also be used with the prime editors.

Description	Sequence
LfCas9 *Lactobacillus fermentum* wild type *GenBank: SNX31424.1* 1	MKEYHIGLDIGTSSIGWAVTDSQFKLMRIKGKTAIGVRLFEEGKTAAERRTFRTTRRRLKR RKWRLHYLDEIFAPHLQEVDENFLRRLKQSNIHPEDPTKNQAFIGKLLFPDLLKKNERGYP TLIKMRDELPVEQRAHYPVMNIYKLREAMINEDRQFDLREVYLAVHHIVKYRGHFLNNA SVDKFKVGRIDFDKSFNVLNEAYEELQNGEGSFTIEPSKVEKIGQLLLDTKMRKLDRQKA VAKLLEVKVADKEETKRNKQIATAMSKLVLGYKADFATVAMANGNEWKIDLSSETSED EIEKFREELSDAQNDILTEITSLFSQIMLNEIVPNGMSISESMMDRYWTHERQLAEVKEYLA TQPASARKEFDQVYNKYIGQAPKERGFDLEKGLKKILSKKENWKEIDELLKAGDFLPKQR

Description	Sequence
	TSANGVIPHQMHQQELDRIIEKQAKYYPWLATENPATGERDRHQAKYELDQLVSFRIPYY VGPLVTPEVQKATSGAKFAWAKRKEDGEITPWNLWDKIDRAESAEAFIKRMTVKDTYLL NEDVLPANSLLYQKYNVLNELNNVRVNGRRLSVGIKQDIYTELFKKKKTVKASDVASLV MAKTRGVNKPSVEGLSDPKKFNSNLATYLDLKSIVGDKVDDNRYQTDLENIIEWRSVFED GEIFADKLTEVEWLTDEQRSALVKKRYKGWGRLSKKLLTGIVDENGQRIIDLMWNTDQN FKEIVDQPVFKEQIDQLNQKAITNDGMTLRERVESVLDDAYTSPQNKKAIWQVVRVVEDI VKAVGNAPKSISIEFARNEGNKGEITRSRRTQLQKLFEDQAHELVKDTSLTEELEKAPDLS DRYYFYFTQGGKDMYTGDPINFDEISTKYDIDHILPQSFVKDNSLDNRVLTSRKENNKKS DQVPAKLYAAKMKPYWNQLLKQGLITQRKFENLTKDVDQNIKYRSLGFVKRQLVETRQ VIKLTANILGSMYQEAGTEIIETRAGLTKQLREEFDLPKVREVNDYHHAVDAYLTTFAGQ YLNRRYPKLRSFFVYGEYMKFKHGSDLKLRNFNFFHELMEGDKSQGKVVDQQTGELITT RDEVAKSFDRLLNMKYMLVSKEVHDRSDQLYGATIVTAKESGKLTSPIEIKKNRLVDLYG AYTNGTSAFMTIIKFTGNKPKYKVIGIPTTSAASLKRAGKPGSESYNQELHRIIKSNPKVKK GFEIVVPHVSYGQLIVDGDCKFTLASPTVQHPATQLVLSKKSLETISSGYKILKDKPAIANE RLIRVFDEVVGQMNRYFTIFDQRSNRQKVADARDKFLSLPTESKYEGAKKVQVGKTEVIT NLLMGLHANATQGDLKVLGLATFGFFQSTTGLSLSEDTMIVYQSPTGLFERRICLKDI (SEQ ID NO: 9)
SaCas9 *Staphylococcus aureus* wild type GenBank: AYD60528.1	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAE ATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGN IVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAL SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDI LRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQE DFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASA QSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAI VDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDN EENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKEL GSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDS IDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIG KATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRM LASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQIS EFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKR YTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 2)
SaCas9 *Staphylococcus aureus*	MGKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRR RRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVH NVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKE AKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCT YFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQI AKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSE DIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVP KKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQK MINEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPF NYEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAK GKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKV KSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQM FEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYSTRKD DKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNP

Description	Sequence
	LYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPY RFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYKNDLI KINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPHIIKTIASKTQSIKKYSTDIL GNLYEVKSKKHPQIIKK (SEQ ID NO: 10)
StCas9 *Streptococcus thermophilus* UniProtKB/Swiss-Prot: G3ECR1.2 Wild type	MLFNKCIIISINLDFSNKEKCMTKPYSIGLDIGTNSVGWAVITDNYKVPSKKMKVLGNTSK KYIKKNLLGVLLFDSGITAEGRRLKRTARRRYTRRRNRILYLQEIFSTEMATLDDAFFQRL DDSFLVPDDKRDSKYPIFGNLVEEKVYHDEFPTIYHLRKYLADSTKKADLRLVYLALAHM IKYRGHFLIEGEFNSKNNDIQKNFQDFLDTYNAIFESDLSLENSKQLEEIVKDKISKLEKKD RILKLFPGEKNSGIFSEFLKLIVGNQADFRKCFNLDEKASLHFSKESYDEDLETLLGYIGDD YSDVFLKAKKLYDAILLSGFLTVTDNETEAPLSSAMIKRYNEHKEDLALLKEYIRNISLKT YNEVFKDDTKNGYAGYIDGKTNQEDFYVYLKNLLAEFEGADYFLEKIDREDFLRKQRTF DNGSIPYQIHLQEMRAILDKQAKFYPFLAKNKERIEKILTFRIPYYVGPLARGNSDFAWSIR KRNEKITPWNFEDVIDKESSAEAFINRMTSFDLYLPEEKVLPKHSLLYETFNVYNELTKVR FIAESMRDYQFLDSKQKKDIVRLYFKDKRKVTDKDIIEYLHAIYGYDGIELKGIEKQFNSSL STYHDLLNIINDKEFLDDSSNEAIIEEIIHTLTIFEDREMIKQRLSKFENIFDKSVLKKLSRRH YTGWGKLSAKLINGIRDEKSGNTILDYLIDDGISNRNFMQLIHDDALSFKKKIQKAQIIGDE DKGNIKEVVKSLPGSPAIKKGILQSIKIVDELVKVMGGRKPESIVVEMARENQYTNQGKSN SQQRLKRLEKSLKELGSKILKENIPAKLSKIDNNALQNDRLYLYYLQNGKDMYTGDDLDI DRLSNYDIDHIIPQAFLKDNSIDNKVLVSSASNRGKSDDFPSLEVVKKRKTFWYQLLKSKL ISQRKFDNLTKAERGGLLPEDKAGFIQRQLVETRQITKHVARLLDEKFNNKKDENNRAVR TVKIITLKSTLVSQFRKDFELYKVREINDFHHAHDAYLNAVIASALLKKYPKLEPEFVYGD YPKYNSFRERKSATEKVYFYSNIMNIFKKSISLADGRVIERPLIEVNEETGESVWNKESDLA TVRRVLSYPQVNVVKKVEEQNHGLDRGKPKGLFNANLSSKPKPNSNENLVGAKEYLDPK KYGGYAGISNSFAVLVKGTIEKGAKKKITNVLEFQGISILDRINYRKDKLNFLLEKGYKDIE LIIELPKYSLFELSDGSRRMLASILSTNNKRGEIHKGNQIFLSQKFVKLLYHAKRISNTINEN HRKYVENHKKEFEELFYYILEFNENYVGAKKNGKLLNSAFQSWQNHSIDELCSSFIGPTGS ERKGLFELTSRGSAADFEFLGVKIPRYRDYTPSSLLKDATLIHQSVTGLYETRIDLAKLGEG (SEQ ID NO: 11)
LcCas9 *Lactobacillus crispatus* NCBI Reference Sequence: WP_133478044.1 Wild type	MKIKNYNLALTPSTSAVGHVEVDDDLNILEPVHHQKAIGVAKFGEGETAEARRLARSARR TTKRRANRINHYFNEIMKPEIDKVDPLMFDRIKQAGLSPLDERKEFRTVIFDRPNIASYYHN QFPTIWHLQKYLMITDEKADIRLIYWALHSLLKHRGHFFNTTPMSQFKPGKLNLKDDMLA LDDYNDLEGLSFAVANSPEIEKVIKDRSMHKKEKIAELKKLIVNDVPDKDLAKRNNKIITQ IVNAIMGNSFHLNFIFDMDLDKLTSKAWSFKLDDPELDTKFDAISGSMTDNQIGIFETLQKI YSAISLLDILNGSSNVVDAKNALYDKHKRDLNLYFKFLNTLPDEIAKTLKAGYTLYIGNRK KDLLAARKLLKVNVAKNFSQDDFYKLINKELKSIDKQGLQTRFSEKVGELVAQNNFLPV QRSSDNVFIPYQLNAITFNKILENQGKYYDFLVKPNPAKKDRKNAPYELSQLMQFTIPYYV GPLVTPEEQVKSGIPKTSRFAWMVRKDNGAITPWNFYDKVDIEATADKFIKRSIAKDSYLL SELVLPKHSLLYEKYEVFNELSNVSLDGKKLSGGVKQILFNEVFKKTNKVNTSRILKALA KHNIPGSKITGLSNPEEFTSSLQTYNAWKKYFPNQIDNFAYQQDLEKMIEWSTVFEDHKIL AKKLDEIEWLDDDQKKFVANTRLRGWGRLSKRLLTGLKDNYGKSIMQRLETTKANFQQI VYKPEFREQIDKISQAAAKNQSLEDILANSYTSPSNRKAIRKTMSVVDEYIKLNHGKEPDK IFLMFQRSEQEKGKQTEARSKQLNRILSQLKADKSANKLFSKQLADEFSNAIKKSKYKLN DKQYFYFQQLGRDALTGEVIDYDELYKYTVLHIIPRSKLTDDSQNNKVLTKYKIVDGSVA LKFGNSYSDALGMPIKAFWTELNRLKLIPKGKLLNLTTDFSTLNKYQRDGYIARQLVETQ QIVKLLATIMQSRFKHTKIIEVRNSQVANIRYQFDYFRIKNLNEYYRGFDAYLAAVVGTYL YKVYPKARRLFVYGQYLKPKKTNQENQDMHLDSEKKSQGFNFLWNLLYGKQDQIFVNG TDVIAFNRKDLITKMNTVYNYKSQKISLAIDYHNGAMFKATLFPRNDRDTAKTRKLIPKK KDYDTDIYGGYTSNVDGYMLLAEIIKRDGNKQYGFYGVPSRLVSELDTLKKTRYTEYEEK LKEIIKPELGVDLKKIKKIKILKNKVPFNQVIIDKGSKFFITSTSYRWNYRQLILSAESQQTL MDLVVDPDFSNHKARKDARKNADERLIKVYEEILYQVKNYMPMFVELHRCYEKLVDAQ KTFKSLKISDKAMVLNQILILLHSNATSPVLEKLGYHTRFTLGKKHNLISENAVLVTQSITG LKENHVSIKQML (SEQ ID NO: 12)

Description	Sequence
PdCas9 *Pedicoccus damnosus* NCBI Reference Sequence: WP_062913273.1 Wild type	MTNEKYSIGLDIGTSSIGFAVVNDNNRVIRVKGKNAIGVRLFDEGKAAADRRSFRTTRRSF RTTRRRLSRRRWRLKLLREIFDAYITPVDEAFFIRLKESNLSPKDSKKQYSGDILFNDRSDK DFYEKYPTIYHLRNALMTEHRKFDVREIYLAIHHIMKFRGHFLNATPANNFKVGRLNLEE KFEELNDIYQRVFPDESIEFRTDNLEQIKEVLLDNKRSRADRQRTLVSDIYQSSEDKDIEKR NKAVATEILKASLGNKAKLNVITNVEVDKEAAKEWSITFDSESIDDDLAKIEGQMTDDGH EIIEVLRSLYSGITLSAIVPENHTLSQSMVAKYDLHKDHLKLFKKLINGMTDTKKAKNLRA AYDGYIDGVKGKVLPQEDFYKQVQVNLDDSAEANEIQTYIDQDIFMPKQRTKANGSIPHQ LQQQELDQIIENQKAYYPWLAELNPNPDKKRQQLAKYKLDELVTFRVPYYVGPMITAKD QKNQSGAEFAWMIRKEPGNITPWNFDQKVDRMATANQFIKRMTTTDTYLLGEDVLPAQS LLYQKFEVLNELNKIRIDHKPISIEQKQQIFNDLFKQFKNVTIKHLQDYLVSQGQYSKRPLI EGLADEKRFNSSLSTYSDLCGIFGAKLVEENDRQEDLEKIIEWSTIFEDKKIYRAKLNDLT WLTDDQKEKLATKRYQGWGRLSRKLLVGLKNSEHRNIMDILWITNENFMQIQAEPDFAK LVTDANKGMLEKTDSQDVINDLYTSPQNKKAIRQILLVVHDIQNAMHGQAPAKIHVEFAR GEERNPRRSVQRQRQVEAAYEKVSNELVSAKVRQEFKEAINNKRDFKDRLFLYFMQGGI DIYTGKQLNIDQLSSYQIDHILPQAFVKDDSLTNRVLTNENQVKADSVPIDIFGKKMLSVW GRMKDQGLISKGKYRNLTMNPENISAHTENGFINRQLVETRQVIKLAVNILADEYGDSTQI ISVKADLSHQMREDFELLKNRDVNDYHHAFDAYLAAFIGNYLLKRYPKLESYFVYGDFK KFTQKETKMRRFNFIYDLKHCDQVVNKETGEILWTKDEDIKYIRHLFAYKKILVSHEVRE KRGALYNQTIYKAKDDKGSGQESKKLIRIKDDKETKIYGGYSGKSLAYMTIVQITKKNKV SYRVIGIPTLALARLNKLENDSTENNGELYKIIKPQFTHYKVDKKNGEIIETTDDFKIVVSK VRFQQLIDDAGQFFMLASDTYKNNAQQLVISNNALKAINNTNITDCPRDDLERLDNLRLD SAFDEIVKKMDKYFSAYDANNFREKIRNSNLIFYQLPVEDQWENNKITELGKRTVLTRILQ GLHANATTTDMSIFKIKTPFGQLRQRSGISLSENAQLIYQSPTGLFERRVQLNKIK (SEQ ID NO: 13)
FnCas9 *Fusobaterium nucleatum* NCBI Reference Sequence: WP_060798984.1	MKKQKFSDYYLGFDIGTNSVGWCVTDLDYNVLRFNKKDMWGSRLFEEAKTAAERRVQ RNSRRRLKRRKWRLNLLEEIFSNEILKIDSNFFRRLKESSLWLEDKSSKEKFTLFNDDNYK DYDFYKQYPTIFHLRNELIKNPEKKDIRLVYLAIHSIFKSRGHFLFEGQNLKEIKNFETLYN NLIAFLEDNGINKIIDKNNIEKLEKIVCDSKKGLKDKEKEFKEIFNSDKQLVAIFKLSVGSSV SLNDLFDTDEYKKGEVEKEKISFREQIYEDDKPIYYSILGEKIELLDIAKTFYDFMVLNNILA DSQYISEAKVKLYEEHKKDLKNLKYIIRKYNKGNYDKLFKDKNENNYSAYIGLNKEKSK KEVIEKSRLKIDDLIKNIKGYLPKVEEIEEKDKAIFNKILNKIELKTILPKQRISDNGTLPYQI HEAELEKILENQSKYYDFLNYEENGIITKDKLLMTFKFRIPYYVGPLNSYHKDKGGNSWIV RKEEGKILPWNFEQKVDIEKSAEEFIKRMTNKCTYLNGEDVIPKDTFLYSEYVILNELNKV QVNDEFLNEENKRKIIDELFKENKKVSEKKFKEYLLVKQIVDGTIELKGVKDSFNSNYISYI RFKDIFGEKLNLDIYKEISEKSILWKCLYGDDKKIFEKKIKNEYGDILTKDEIKKINTFKFNN WGRLSEKLLTGIEFINLETGECYSSVMDALRRTNYNLMELLSSKFTLQESINNENKEMNEA SYRDLIEESYVSPSLKRAIFQTLKIYEEIRKITGRVPKKVFIEMARGGDESMKNKKIPARQE QLKKLYDSCGNDIANFSIDIKEMKNSLISYDNNSLRQKKLYLYYLQFGKCMYTGREIDLD RLLQNNDTYDIDHIYPRSKVIKDDSFDNLVLVLKNENAEKSNEYPVKKEIQEKMKSFWRF LKEKNFISDEKYKRLTGKDDFELRGFMARQLVNVRQTTKEVGKILQQIEPEIKIVYSKAEI ASSFREMFDFIKVRELNDTHHAKDAYLNIVAGNVYNTKFTEKPYRYLQEIKENYDVKKIY NYDIKNAWDKENSLEIVKKNMEKNTVNITRFIKEKKGQLFDLNPIKKGETSNEIISIKPKVY NGKDDKLNEKYGYYKSLNPAYFLYVEHKEKNKRIKSFERVNLVDVNNIKDEKSLVKYLI ENKKLVEPRVIKKVYKRQVILINDYPYSIVTLDSNKLMDFENLKPLFLENKYEKILKNVIKF LEDNQGKSEENYKFIYLKKKDRYEKNETLESVKDRYNLEFNEMYDKFLEKLDSKDYKNY MNNKKYQELLDVKEKFIKLNLFDKAFTLKSFLDLFNRKTMADFSKVGLTKYLGKIQKISS NVLSKNELYLLEESVTGLFVKKIKL (SEQ ID NO: 14)
EcCas9 *Enterococcus cecorum* NCBI Reference Sequence: WP_047338501.1	RRKQRIQILQELLGEEVLKTDPGFFHRMKESRYVVEDKRTLDGKQVELPYALFVDKDYTD KEYYKQFPTINHLIVYLMTTSDTPDIRLVYLALHYYMKNRGNFLHSGDINNVKDINDILEQ LDNVLETFLDGWNLKLKSYVEDIKNIYNRDLGRGERKKAFVNTLGAKTKAEKAFCSLISG GSTNLAELFDDSSLKEIETPKIEFASSSLEDKIDGIQEALEDRFAVIEAAKRLYDWKTLTDIL GDSSSLAEARVNSYQMHHEQLLELKSLVKEYLDRKVFQEVFVSLNVANNYPAYIGHTKI NGKKKELEVKRTKRNDFYSYVKKQVIEPIKKKVSDEAVLTKLSEIESLIEVDKYLPLQVNS DNGVIPYQVKLNELTRIFDNLENRIPVLRENRDKIIKTFKFRIPYYVGSLNGVVKNGKCTN

Description	Sequence
Wild type	WMVRKEEGKIYPWNFEDKVDLEASAEQFIRRMTNKCTYLVNEDVLPKYSLLYSKYLVLS ELNNLRIDGRPLDVKIKQDIYENVFKKNRKVTLKKIKKYLLKEGIITDDDELSGLADDVKS SLTAYRDFKEKLGHLDLSEAQMENIILNITLFGDDKKLLKKRLAALYPFIDDKSLNRIATLN YRDWGRLSERFLSGITSVDQETGELRTIIQCMYETQANLMQLLAEPYHFVEAIEKENPKVD LESISYRIVNDLYVSPAVKRQIWQTLLVIKDIKQVMKHDPERIFIEMARENQESKKTKSRK QVLSEVYKKAKEYEHLFEKLNSLTEEQLRSKKIYLYFTQLGKCMYSGEPIDFENLVSANS NYDIDHIYPQSKTIDDSFNNIVLVKKSLNAYKSNHYPIDKNIRDNEKVKTLWNTLVSKGLI TKEKYERLIRSTPFSDEELAGFIARQLVETRQSTKAVAEILSNWFPESEIVYSKAKNVSNFR QDFEILKVRELNDCHHAHDAYLNIVVGNAYHTKFTNSPYRFIKNKANQEYNLRKLLQKV NKIESNGVVAWVGQSENNPGTIATVKKVIRRNTVLISRMVKEVDGQLFDLTLMKKGKGQ VPIKSSDERLTDISKYGGYNKATGAYFTFVKSKKRGKVVRSFEYVPLHLSKQFENNNELL KEYIEKDRGLTDVEILIPKVLINSLFRYNGSLVRITGRGDTRLLLVHEQPLYVSNSFVQQLK SVSSYKLKKSENDNAKLTKTATEKLSNIDELYDGLLRKLDLPIYSYWFSSIKEYLVESRTK YIKLSIEEKALVIFEILHLFQSDAQVPNLKILGLSTKPSRIRIQKNLKDTDKMSIIHQSPSGIFE HEIELTSL (SEQ ID NO: 15)
AhCas9 *Anaerostipes hadrus* NCBI Reference Sequence: WP_044924278.1 Wild type	MQNGFLGITVSSEQVGWAVTNPKYELERASRKDLWGVRLFDKAETAEDRRMFRTNRRL NQRKKNRIHYLRDIFHEEVNQKDPNFFQQLDESNFCEDDRTVEFNFDTNLYKNQFPTVYH LRKYLMETKDKPDIRLVYLAFSKFMKNRGHFLYKGNLGEVMDFENSMKGFCESLEKFNI DFPTLSDEQVKEVRDILCDHKIAKTVKKKNIITITKVKSKTAKAWIGLFCGCSVPVKVLFQ DIDEEIVTDPEKISFEDASYDDYIANIEKGVGIYYEAIVSAKMLFDWSILNEILGDHQLLSDA MIAEYNKHHDDLKRLQKIIKGTGSRELYQDIFINDVSGNYVCYVGHAKTMSSADQKQFY TFLKNRLKNVNGISSEDAEWIDTEIKNGTLLPKQTKRDNSVIPHQLQLREFELILDNMQEM YPFLKENREKLLKIFNFVIPYYVGPLKGVVRKGESTNWMVPKKDGVIHPWNFDEMVDKE ASAECFISRMTGNCSYLFNEKVLPKNSLLYETFEVLNELNPLKINGEPISVELKQRIYEQLF LTGKKVTKKSLTKYLIKNGYDKDIELSGIDNEFHSNLKSHIDFEDYDNLSDEEVEQIILRITV FEDKQLLKDYLNREFVKLSEDERKQICSLSYKGWGNLSEMLLNGITVTDSNGVEVSVMD MLWNTNLNLMQILSKKYGYKAEIEHYNKEHEKTIYNREDLMDYLNIPPAQRRKVNQLITI VKSLKKTYGVPNKIFFKISRERQDDPKRTSSRKEQLKYLYKSLKSEDEKHLMKELDELND HELSNDKVYLYFLQKGRCIYSGKKLNLSRLRKSNYQNDIDYIYPLSAVNDRSMNNKVLTG IQENRADKYTYFPVDSEIQKKMKGFWMELVLQGFMTKEKYFRLSRENDFSKSELVSFIER EISDNQQSGRMIASVLQYYFPESKIVFVKEKLISSFKRDFHLISSYGHNHLQAAKDAYITIV VGNVYHTKFTMDPAIYFKNHKRKDYDLNRLFLENISRDGQIAWESGPYGSIQTVRKEYAQ NHIAVTKRVVEVKGGLFKQMPLKKGHGEYPLKTNDPRFGNIAQYGGYTNVTGSYFVLVE SMEKGKKRISLEYVPVYLHERLEDDPGHKLLKEYLVDHRKLNHPKILLAKVRKNSLLKID GFYYRLNGRSGNALILTNAVELIMDDWQTKTANKISGYMKRRAIDKKARVYQNEFHIQE LEQLYDFYLDKLKNGVYKNRKNNQAELIHNEKEQFMELKTEDQCVLLTEIKKLFVCSPM QADLTLIGGSKHTGMIAMSSNVTKADFAVIAEDPLGLRNKVIYSHKGEK (SEQ ID NO: 16)
KvCas9 *Kandleria vitulina* NCBI Reference Sequence: WP_031589969.1 Wild type	MSQNNNKIYNIGLDIGDASVGWAVVDEHYNLLKRHGKHMWGSRLFTQANTAVERRSSR STRRRYNKRRERIRLLREIMEDMVLDVDPTFFIRLANVSFLDQEDKKDYLKENYHSNYNL FIDKDFNDKTYYDKYPTIYHLRKHLCESKEKEDPRLIYLALHHIVKYRGNFLYEGQKFSM DVSNIEDKMIDVLRQFNEINLFEYVEDRKKIDEVLNVLKEPLSKKHKAEKAFALFDTTKD NKAAYKELCAALAGNKFNVTKMLKEAELHDEDEKDISFKFSDATFDDAFVEKQPLLGDC VEFIDLLHDIYSWVELQNILGSAHTSEPSISAAMIQRYEDHKNDLKLLKDVIRKYLPKKYF EVFRDEKSKKNNYCNYINHPSKTPVDEFYKYIKKLIEKIDDPDVKTILNKIELESFMLKQNS RTNGAVPYQMQLDELNKILENQSVYYSDLKDNEDKIRSILTFRIPYYFGPLNITKDRQFDW IIKKEGKENERILPWNANEIVDVDKTADEFIKRMRNFCTYFPDEPVMAKNSLTVSKYEVL NEINKLRINDHLIKRDMKDKMLHTLFMDHKSISANAMKKWLVKNQYFSNTDDIKIEGFQ KENACSTSLTPWIDFTKIFGKINESNYDFIEKIIYDVTVFEDKKILRRRLKKEYDLDEEKIKK ILKLKYSGWSRLSKKLLSGIKTKYKDSTRTPETVLEVMERTNMNLMQVINDEKLGFKKTI DDANSTSVSGKFSYAEVQELAGSPAIKRGIWQALLIVDEIKKIMKHEPAHVYIEFARNEDE KERKDSFVNQMLKLYKDYDFDEDETEKEANKHLKGEDAKSKIRSERLKLYYTQMGKCMY TGKSLDIDRLDTYQVDHIVPQSLLKDDSIDNKVLVLSSENQRKLDDLVIPSSIRNKMYGFW EKLFNNKIISPKKFYSLIKTEFNEKDQERFINRQIVETRQITKHVAQIIDNHYENTKVVTVRA DLSHQFRERYHIYKNRDINDFHHAHDAYIATILGTYIGHRFESLDAKYIYGEYKRIFRNQK

Description	Sequence
	NKGKEMKKNNDGFILNSMRNIYADKDTGEIVWDPNYIDRIKKCFYYKDCFVTKKLEENN GTFFNVTVLPNDTNSDKDNTLATVPVNKYRSNVNKYGGFSGVNSFIVAIKGKKKKGKKV IEVNKLTGIPLMYKNADEEIKINYLKQAEDLEEVQIGKEILKNQLIEKDGGLYYIVAPTEIIN AKQLILNESQTKLVCEIYKAMKYKNYDNLDSEKIIDLYRLLINKMELYYPEYRKQLVKKF EDRYEQLKVISIEEKCNIIKQILATLHCNSSIGKIMYSDFKISTTIGRLNGRTISLDDISFIAESP TGMYSKKYKL (SEQ ID NO: 17)
EfCas9 *Enterococcus faecalis* NCBI Reference Sequence: WP_016631044.1 Wild type	MRLFEEGHTAEDRRLKRTARRRISRRRNRLRYLQAFFEEAMTDLDENFFARLQESFLVPE DKKWHRHPIFAKLEDEVAYHETYPTIYHLRKKLADSSEQADLRLIYLALAHIVKYRGHFLI EGKLSTENTSVKDQFQQFMVIYNQTFVNGESRLVSAPLPESVLIEEELTEKASRTKKSEKV LQQFPQEKANGLFGQFLKLMVGNKADFKKVFGLEEEAKITYASESYEEDLEGILAKVGDE YSDVFLAAKNVYDAVELSTILADSDKKSHAKLSSSMIVRFTEHQEDLKKFKRFIRENCPDE YDNLFKNEQKDGYAGYIAHAGKVSQLKFYQYVKKIIQDIAGAEYFLEKIAQENFLRKQRT FDNGVIPHQIHLAELQAIIHRQAAYYPFLKENQEKIEQLVTFRIPYYVGPLSKGDASTFAWL KRQSEEPIRPWNLQETVDLDQSATAFIERMTNFDTYLPSEKVLPKHSLLYEKFMVFNELTK ISYTDDRGIKANFSGKEKEKIFDYLFKTRRKVKKKDIIQFYRNEYNTEIVTLSGLEEDQFNA SFSTYQDLLKCGLTRAELDHPDNAEKLEDIIKILTIFEDRQRIRTQLSTFKGQFSAEVLKKLE RKHYTGWGRLSKKLINGIYDKESGKTILDYLVKDDGVSKHYNRNFMQLINDSQLSFKNAI QKAQSSEHEETLSETVNELAGSPAIKKGIYQSLKIVDELVAIMGYAPKRIVVEMARENQTT STGKRRSIQRLKIVEKAMAEIGSNLLKEQPTTNEQLRDTRLFLYYMQNGKDMYTGDELSL HRLSHYDIDHIIPQSFMKDDSLDNLVLVGSTENRGKSDDVPSKEVVKDMKAYWEKLYAA GLISQRKFQRLTKGEQGGLTLEDKAHFIQRQLVETRQITKNVAGILDQRYNAKSKEKKVQI ITLKASLTSQFRSIFGLYKVREVNDYHHGQDAYLNCVVATTLLKVYPNLAPEFVYGEYPK FQTFKENKATAKAIIYTNLLRFFTEDEPRFTKDGEILWSNSYLKTIKKELNYHQMNIVKKV EVQKGGFSKESIKPKGPSNKLIPVKNGLDPQKYGGFDSPVVAYTVLFTHEKGKKPLIKQEI LGITIMEKTRFEQNPILFLEEKGFLRPRVLMKLPKYTLYEFPEGRRRLLASAKEAQKGNQM VLPEHLLTLLYHAKQCLLPNQSESLAYVEQHQPEFQEILERVVDFAEVHTLAKSKVQQIV KLFEANQTADVKEIAASFIQLMQFNAMGAPSTFKFFQKDIERARYTSIKEIFDATIIYQSPTG LYETRRKVVD (SEQ ID NO: 18)
Staphylococcus aureus Cas9	KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRH RIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVN EVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQ LLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFP EELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKKPTLKQIAK EILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELLDQIAKILTIYQSSEDIQ EELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKK VDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMI NEMQKRNRQTNERIEEIIRTTGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFN YEVDHIIPRSVSFDNSFNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKG KGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVK SINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMF EEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDD KGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPL YKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYR FDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIK INGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILG NLYEVKSKKHPQIIKKG (SEQ ID NO: 19)
Geobacillus thermodenitrificans Cas9	MKYKIGLDIGITSIGWAVINLDIPRIEDLGVRIFDRAENPKTGESLALPRRLARSARRRLRRR KHRLERIRRLFVREGILTKEELNKLFEKKHEIDVWQLRVEALDRKLNNDELARILLHLAKR RGFRSNRKSERTNKENSTMLKHIEENQSILSSYRTVAEMVVKDPKFSLHKRNKEDNYTNT VARDDLEREIKLIFAKQREYGNIVCTEAFEHEYISIWASQRPFASKDDIEKKVGFCTFEPKE KRAPKATYTFQSFTVWEHINKLRLVSPGGIRALTDDERRLIYKQAFHKNKITFHDVRTLLN LPDDTRFKGLLYDRNTTLKENEKVRFLELGAYHKIRKAIDSVYGKGAAKSFRPIDFDTFG YALTMFKDDTDIRSYLRNEYEQNGKRMENLADKVYDEELIEELLNLSFSKFGHLSLKALR NILPYMEQGEVYSTACERAGYTFTGPKKKQKTVLLPNIPPIANPVVMRALTQARKVVNAII

Description	Sequence
	KKYGSPVSIHIELARELSQSFDERRKMQKEQEGNRKKNETAIRQLVEYGLTLNPTGLDIVK FKLWSEQNGKCAYSLQPIEIERLLEPGYTEVDHVIPYSRSLDDSYTNKVLVLTKENREKGN RTPAEYLGLGSERWQQFETFVLTNKQFSKKKRDRLLRLHYDENEENEFKNRNLNDTRYIS RFLANFIREHLKFADSDDKQKVYTVNGRITAHLRSRWNFNKNREESNLHHAVDAAIVAC TTPSDIARVTAFYQRREQNKELSKKTDPQFPQPWPHFADELQARLSKNPKESIKALNLGN YDNEKLESLQPVFVSRMPKRSITGAAHQETLRRYIGIDERSGKIQTVVKKKLSEIQLDKTG HFPMYGKESDPRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGELGPIIRTIKIIDTTNQVIP LNDGKTVAYNSNIVRVDVFEKDGKYYCVPIYTIDMMKGILPNKAIEPNKPYSEWKEMTE DYTFRFSLYPNDLIRIEFPREKTIKTAVGEEIKIKDLFAYYQTIDSSNGGLSLVSHDNNFSLR SIGSRTLKRFEKYQVDVLGNIYKVRGEKRVGVASSSHSKAGETIRPL (SEQ ID NO: 20)
ScCas9 *S. canis* 1375 AA 159.2 kDa	MEKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTNRKSIKKNLMGALLFDSGETAE ATRLKRTARRRYTRRKNRIRYLQEIFANEMAKLDDSFFQRLEESFLVEEDKKNERHPIFGN LADEVAYHRNYPTIYHLRKKLADSPEKADLRLIYLALAHIIKFRGHFLIEGKLNAENSDVA KLFYQLIQTYNQLFEESPLDEIEVDAKGILSARLSKSKRLEKLIAVFPNEKKNGLFGNIIALA LGLTPNFKSNFDLTEDAKLQLSKDTYDDDLDELLGQIGDQYADLFSAAKNLSDAILLSDIL RSNSEVTKAPLSASMVKRYDEHHQDLALLKTLVRQQFPEKYAEIFKDDTKNGYAGYVDG GIKHRKRTTKLATQEEFYKFIKPILEKMDGAEELLAKLNREDLLRKQRTFDNGSIPHQIHL KELHAILRRQEEFYPFLKENREKIEKILTFRIPYYVGPLARGNSRFAWLTRKSEEAITPWNF EEVVDKGASAQSFIERMTNFDEQLPNEKVLPKHSLLYEYFTVYNELTKVKYVTERMRKP EFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEIIGVEDRFNASLGTYHDLL KIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRHYTG WGRLSRKMINGIRDKQSGKTILDFLKSDGFSNRNFMQLIHDDSLTFKEEIEKAQVSGQGDS LHEQIANLAGSPAIKKGILQTVKIVDELVKVMGHKPENIVIEMARENQTTTKGLQQSRERK KIEEGIKELESQILKENPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFIKDDSIDNKVLTRSVENRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLT KAERGGLSEADKAGFIKRQLVETRQITKHVARILDSRMNTKRDKNDKPIREVKVITLKSKL VSDFRKDFQLYKVRDINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRK MIAKSEQEIGKATAKRFFYSNIMNFFKTEVKLANGEIRKRPLIETNGETGEVVWNKEKDFA TVRKVLAMPQVNIVKKTEVQTGGFSKESILSKRESAKLIPRKKGWDTRKYGGFGSPTVAY SILVVAKVEKGKAKKLKSVKVLVGITIMEKGSYEKDPIGFLEAKGYKDIKKELIFKLPKYS LFELENGRRRMLASATELQKANELVLPQHLVRLLYYTQNISATTGSNNLGYIEQHREEFK EIFEKIIDFSEKYILKNKVNSNLKSSFDEQFAVSDSILLSNSFVSLLKYTSFGASGGFTFLDLD VKQGRLRYQTVTEVLDATLIYQSITGLYETRTDLSQLGGD (SEQ ID NO: 21)

[334] The prime editors utilized in the methods and compositions described herein may include any of the above Cas9 ortholog sequences, or any variants thereof having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto.

[335] The napDNAbp may include any suitable homologs and/or orthologs or naturally occurring enzymes, such as, Cas9. Cas9 homologs and/or orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. Preferably, the Cas moiety is configured (e.g., mutagenized, recombinantly engineered, or otherwise obtained from nature) as a nickase, i.e., capable of cleaving only a single strand of the target double-stranded DNA. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the organisms and loci disclosed in Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) *RNA Biology*

10:5, 726-737; the entire contents of which are incorporated herein by reference. In some embodiments, a Cas9 nuclease has an inactive (e.g., an inactivated) DNA cleavage domain; that is, the Cas9 is a nickase. In some embodiments, the Cas9 protein comprises an amino acid sequence that is at least 80% identical to the amino acid sequence of a Cas9 protein as provided by any one of the variants of Table 3. In some embodiments, the Cas9 protein comprises an amino acid sequence that is at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence of a Cas9 protein as provided by any one of the Cas9 orthologs in the above tables.

C. Dead Cas9 variant

[336] In certain embodiments, the prime editors utilized in the methods and compositions described herein may include a dead Cas9, e.g., dead SpCas9, which has no nuclease activity due to one or more mutations that inactive both nuclease domains of Cas9, namely the RuvC domain (which cleaves the non-protospacer DNA strand) and HNH domain (which cleaves the protospacer DNA strand). The nuclease inactivation may be due to one or mutations that result in one or more substitutions and/or deletions in the amino acid sequence of the encoded protein, or any variants thereof having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto.

[337] As used herein, the term "dCas9" refers to a nuclease-inactive Cas9 or nuclease-dead Cas9, or a functional fragment thereof, and embraces any naturally occurring dCas9 from any organism, any naturally-occurring dCas9 equivalent or functional fragment thereof, any engineered dCas9 variant or functional fragment thereof, any dCas9 homolog, ortholog, or paralog from any organism, and any mutant or variant of a dCas9, naturally-occurring or engineered. The term dCas9 is not meant to be particularly limiting and may be referred to as a "dCas9 or equivalent." Exemplary dCas9 proteins and method for making dCas9 proteins are further described herein and/or are described in the art and are incorporated herein by reference.

[338] In other embodiments, dCas9 corresponds to, or comprises in part or in whole, a Cas9 amino acid sequence having one or more mutations that inactivate the Cas9 nuclease activity. In other embodiments, Cas9 variants having mutations other than D10A and H840A are provided which may result in the full or partial inactivation of the endogenous Cas9 nuclease activity (e.g., nCas9 or dCas9, respectively). Such mutations, by way of example, include other amino acid substitutions at D10 and H840, or other substitutions within the nuclease domains of Cas9 (e.g.,

substitutions in the HNH nuclease subdomain and/or the RuvC1 subdomain) with reference to a wild type sequence such as Cas9 from *Streptococcus pyogenes* (NCBI Reference Sequence: NC_017053.1). In some embodiments, variants or homologues of Cas9 (e.g., variants of Cas9 from *Streptococcus pyogenes* (NCBI Reference Sequence: NC_017053.1 (SEQ ID NO: 4))) are provided which are at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to NCBI Reference Sequence: NC_017053.1. In some embodiments, variants of dCas9 (e.g., variants of NCBI Reference Sequence: NC_017053.1 (SEQ ID NO: 4)) are provided having amino acid sequences which are shorter, or longer than NC_017053.1 (SEQ ID NO: 4) by about 5 amino acids, by about 10 amino acids, by about 15 amino acids, by about 20 amino acids, by about 25 amino acids, by about 30 amino acids, by about 40 amino acids, by about 50 amino acids, by about 75 amino acids, by about 100 amino acids or more.

[339] In one embodiment, the dead Cas9 may be based on the canonical SpCas9 sequence of Q99ZW2 and may have the following sequence, which comprises a D10X and an H810X, wherein X may be any amino acid, substitutions (underlined and bolded), or a variant be variant of SEQ ID NO: 40 having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto.

[340] In one embodiment, the dead Cas9 may be based on the canonical SpCas9 sequence of Q99ZW2 and may have the following sequence, which comprises a D10A and an H810A substitutions (underlined and bolded), or be a variant of SEQ ID NO: 23 having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto.

Description	Sequence	SEQ ID NO:
dead Cas9 or dCas9 *Streptococcus pyogenes* Q99ZW2 Cas9 with D10**X** and H810**X** Where "X" is any amino acid	MDKKYSIGL**X**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS	22

Description	Sequence	SEQ ID NO:
	GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD**X**I VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGD	
dead Cas9 or dCas9 *Streptococcus pyogenes* Q99ZW2 Cas9 with D10**A** and H810**A**	MDKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD**A**I VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGD	23

D. Cas9 nickase variant

[341] In one embodiment, the prime editors utilized in the methods and compositions described herein comprise a Cas9 nickase. The term "Cas9 nickase" or "nCas9" refers to a variant of Cas9 which is capable of introducing a single-strand break in a double strand DNA molecule target. In some embodiments, the Cas9 nickase comprises only a single functioning nuclease domain. The wild type Cas9 (e.g., the canonical SpCas9) comprises two separate nuclease domains, namely, the RuvC domain (which cleaves the non-protospacer DNA strand) and HNH domain (which cleaves the protospacer DNA strand). In one embodiment, the Cas9 nickase comprises a

mutation in the RuvC domain which inactivates the RuvC nuclease activity. For example, mutations in aspartate (D) 10, histidine (H) 983, aspartate (D) 986, or glutamate (E) 762, have been reported as loss-of-function mutations of the RuvC nuclease domain and the creation of a functional Cas9 nickase (e.g., Nishimasu *et al.*, "Crystal structure of Cas9 in complex with guide RNA and target DNA," *Cell* 156(5), 935–949, which is incorporated herein by reference). Thus, nickase mutations in the RuvC domain could include D10X, H983X, D986X, or E762X, wherein X is any amino acid other than the wild type amino acid. In certain embodiments, the nickase could be D10A, of H983A, D986A, or E762A, or a combination thereof.

[342] In various embodiments, the Cas9 nickase can have a mutation in the RuvC nuclease domain and have one of the following amino acid sequences, or a variant thereof having an amino acid sequence that has at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto.

Description	Sequence	SEQ ID NO:
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with D10X, wherein X is any alternate amino acid	MDKKYSIGLXIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	24
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with E762X,	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL	25

Description	Sequence	SEQ ID NO:
wherein X is any alternate amino acid	TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIXMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVD QELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDA YLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSN IMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKK TEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELEN GRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNL GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with H983X, wherein X is any alternate amino acid	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHXAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	26
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with D986X, wherein X is any alternate amino acid	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD	27

Description	Sequence	SEQ ID NO:
	KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAH**X**AYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with D10**A**	MDKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	28
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with E762A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL	29

Description	Sequence	SEQ ID NO:
	SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIAMARENQTTQ KGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVD QELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKM KNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQ ILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDA YLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSN IMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKK TEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEK GKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELEN GRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNL GAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with H983A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHAAHDAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	30
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with D986A	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEED KKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRR LENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLD NLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTE ELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFD KNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLF KTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLD NEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRL SRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQG DSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ	31

Description	Sequence	SEQ ID NO:
	ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQIL DSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHAAYL NAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIM NFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGK SKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	

[343] In another embodiment, the Cas9 nickase comprises a mutation in the HNH domain which inactivates the HNH nuclease activity. For example, mutations in histidine (H) 840 or asparagine (R) 863 have been reported as loss-of-function mutations of the HNH nuclease domain and the creation of a functional Cas9 nickase (e.g., Nishimasu *et al.*, "Crystal structure of Cas9 in complex with guide RNA and target DNA," *Cell* 156(5), 935–949, which is incorporated herein by reference). Thus, nickase mutations in the HNH domain could include H840X and R863X, wherein X is any amino acid other than the wild type amino acid. In certain embodiments, the nickase could be H840A or R863A or a combination thereof.

[344] In various embodiments, the Cas9 nickase can have a mutation in the HNH nuclease domain and have one of the following amino acid sequences, or a variant thereof having an amino acid sequence that has at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto.

Description	Sequence	SEQ ID NO:
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with H840X, wherein X is any alternate amino acid	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDXI VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI	32

Description	Sequence	SEQ ID NO:
	ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGD	
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with H840<u>A</u>	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD<u>A</u>I VPQSFLKDDSIDNKVLTRSDKN<u>R</u>GKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGD	33
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with R863X, wherein X is any alternate amino acid	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKN<u>X</u>GKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS	34

Description	Sequence	SEQ ID NO:
	KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGD	
Cas9 nickase *Streptococcus pyogenes* Q99ZW2 Cas9 with R863<u>A</u>	MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSG ETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKK HERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHF LIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQI GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLN REDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIV LTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQS GKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSP AIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHI VPQSFLKDDSIDNKVLTRSDKN<u>A</u>GKSDNVPSEEVVKKMKNYWRQLLNAKLITQR KFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLES EFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLI ETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKL IARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANL DKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLD ATLIHQSITGLYETRIDLSQLGGD	35

[345] In some embodiments, the N-terminal methionine is removed from a Cas9 nickase, or from any Cas9 variant, ortholog, or equivalent disclosed or contemplated herein. For example, methionine-minus Cas9 nickases include the following sequences, or a variant thereof having an amino acid sequence that has at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity thereto.

Description	Sequence
Cas9 nickase (Met minus) *Streptococcus pyogenes* Q99ZW2 Cas9 with H840<u>X</u>, wherein X is any alternate amino acid	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAL SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSD ILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFL DNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLI NGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD<u>X</u>IVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE

	RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVS DFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQH KHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 36)
Cas9 nickase (Met minus) *Streptococcus pyogenes* Q99ZW2 Cas9 with H840A	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAL SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSD ILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFL DNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLI NGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVD**A**IVPQSF LKDDSIDNKVLTRSDKN**R**GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVS DFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQH KHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 37)
Cas9 nickase (Met minus) *Streptococcus pyogenes* Q99ZW2 Cas9 with R863X, wherein X is any alternate amino acid	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAL SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSD ILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFL DNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLI NGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKN**X**GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVS DFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQH KHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD(SEQ ID NO: 38)

Cas9 nickase (Met minus) *Streptococcus pyogenes* Q99ZW2 Cas9 with R863A	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEA TRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNI VDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIAL SLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSD ILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDG GASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQ EDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFL DNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLI NGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF LKDDSIDNKVLTRSDKNAGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVS DFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATV RKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYS VLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQH KHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAF KYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 39)

E. Other Cas9 variants

[346] Besides dead Cas9 and Cas9 nickase variants, the Cas9 proteins used herein may also include other "Cas9 variants" having at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to any reference Cas9 protein, including any wild type Cas9, or mutant Cas9 (e.g., a dead Cas9 or Cas9 nickase), or fragment Cas9, or circular permutant Cas9, or other variant of Cas9 disclosed herein or known in the art. In some embodiments, a Cas9 variant may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 or more amino acid changes compared to a reference Cas9. In some embodiments, the Cas9 variant comprises a fragment of a reference Cas9 (*e.g.*, a gRNA binding domain or a DNA-cleavage domain), such that the fragment is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to the corresponding fragment of wild type Cas9. In some embodiments, the fragment is at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at

least 75%, at least 80%, at least 85%, at least 90%, at least 95% identical, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% of the amino acid length of a corresponding wild type Cas9 (*e.g.*, SEQ ID NO: 2).

[347] In some embodiments, the disclosure also may utilize Cas9 fragments that retain their functionality and that are fragments of any herein disclosed Cas9 protein. In some embodiments, the Cas9 fragment is at least 100 amino acids in length. In some embodiments, the fragment is at least 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, or at least 1300 amino acids in length.

[348] In various embodiments, the prime editors utilized in the methods and compositions disclosed herein may comprise one of the Cas9 variants described as follows, or a Cas9 variant thereof having at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to any reference Cas9 variants.

F. Small-sized Cas9 variants

[349] In some embodiments, the prime editors utilized in the methods and compositions contemplated herein can include a Cas9 protein that is of smaller molecular weight than the canonical SpCas9 sequence. In some embodiments, the smaller-sized Cas9 variants may facilitate delivery to cells, e.g., by an expression vector, nanoparticle, or other means of delivery. In certain embodiments, the smaller-sized Cas9 variants can include enzymes categorized as type II enzymes of the Class 2 CRISPR-Cas systems. In some embodiments, the smaller-sized Cas9 variants can include enzymes categorized as type V enzymes of the Class 2 CRISPR-Cas systems. In other embodiments, the smaller-sized Cas9 variants can include enzymes categorized as type VI enzymes of the Class 2 CRISPR-Cas systems.

[350] The canonical SpCas9 protein is 1368 amino acids in length and has a predicted molecular weight of 158 kilodaltons. The term "small-sized Cas9 variant", as used herein, refers to any Cas9 variant—naturally occurring, engineered, or otherwise—that is less than at least 1300 amino acids, or at least less than 1290 amino acids, or than less than 1280 amino acids, or less than 1270 amino acid, or less than 1260 amino acid, or less than 1250 amino acids, or less than 1240 amino acids, or less than 1230 amino acids, or less than 1220 amino acids, or less than 1210 amino acids, or less than 1200 amino acids, or less than 1190 amino acids, or less than

1180 amino acids, or less than 1170 amino acids, or less than 1160 amino acids, or less than 1150 amino acids, or less than 1140 amino acids, or less than 1130 amino acids, or less than 1120 amino acids, or less than 1110 amino acids, or less than 1100 amino acids, or less than 1050 amino acids, or less than 1000 amino acids, or less than 950 amino acids, or less than 900 amino acids, or less than 850 amino acids, or less than 800 amino acids, or less than 750 amino acids, or less than 700 amino acids, or less than 650 amino acids, or less than 600 amino acids, or less than 550 amino acids, or less than 500 amino acids, but at least larger than about 400 amino acids and retaining the required functions of the Cas9 protein. The Cas9 variants can include those categorized as type II, type V, or type VI enzymes of the Class 2 CRISPR-Cas system.

[351] In various embodiments, the prime editors utilized in the methods and compositions disclosed herein may comprise one of the small-sized Cas9 variants described as follows, or a Cas9 variant thereof having at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to any reference small-sized Cas9 protein.

Description	Sequence	SEQ ID NO:
SaCas9 *Staphylococcus aureus* 1053 AA 123 kDa	MGKRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARR LKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLH LAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRG SINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFG WKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLE YYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIK DITARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTH NLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVK RSFIQSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRT TGKENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNS FNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKE YLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGF TSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFE EKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRKLINDTLYST RKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIME QYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNS RNKVVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLK KISNQAEFIASFYKNDLIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDK RPPHIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKK	10
NmeCas9 *N. meningitidis* 1083 AA 124.5 kDa	MAAFKPNSINYILGLDIGIASVGWAMVEIDEEENPIRLIDLGVRVFERAEVPKTGDS LAMARRLARSVRRLTRRRAHRLLRTRRLLKREGVLQAANFDENGLIKSLPNTPWQ LRAAALDRKLTPLEWSAVLLHLIKHRGYLSQRKNEGETADKELGALLKGVAGNA HALQTGDFRTPAELALNKFEKESGHIRNQRSDYSHTFSRKDLQAELILLFEKQKEFG NPHVSGGLKEGIETLLMTQRPALSGDAVQKMLGHCTFEPAEPKAAKNTYTAERFI WLTKLNNLRILEQGSERPLTDTERATLMDEPYRKSKLTYAQARKLLGLEDTAFFK GLRYGKDNAEASTLMEMKAYHAISRALEKEGLKDKKSPLNLSPELQDEIGTAFSLF	40

Description	Sequence	SEQ ID NO:
	KTDEDITGRLKDRIQPEILEALLKHISFDKFVQISLKALRRIVPLMEQGKRYDEACAE IYGDHYGKKNTEEKIYLPPIPADEIRNPVVLRALSQARKVINGVVRRYGSPARIHIET AREVGKSFKDRKEIEKRQEENRKDREKAAAKFREYFPNFVGEPKSKDILKLRLYEQ QHGKCLYSGKEINLGRLNEKGYVEIDAALPFSRTWDDSFNNKVLVLGSENQNKGN QTPYEYFNGKDNSREWQEFKARVETSRFPRSKKQRILLQKFDEDGFKERNLNDTR YVNRFLCQFVADRMRLTGKGKKRVFASNGQITNLLRGFWGLRKVRAENDRHHAL DAVVVACSTVAMQQKITRFVRYKEMNAFDGKTIDKETGEVLHQKTHFPQPWEFF AQEVMIRVFGKPDGKPEFEEADTLEKLRTLLAEKLSSRPEAVHEYVTPLFVSRAPN RKMSGQGHMETVKSAKRLDEGVSVLRVPLTQLKLKDLEKMVNREREPKLYEALK ARLEAHKDDPAKAFAEPFYKYDKAGNRTQQVKAVRVEQVQKTGVWVRNHNGIA DNATMVRVDVFEKGDKYYLVPIYSWQVAKGILPDRAVVQGKDEEDWQLIDDSFN FKFSLHPNDLVEVITKKARMFGYFASCHRGTGNINIRIHDLDHKIGKNGILEGIGVK TALSFQKYQIDELGKEIRPCRLKKRPPVR	
CjCas9 *C. jejuni* 984 AA 114.9 kDa	MARILAFDIGISSIGWAFSENDELKDCGVRIFTKVENPKTGESLALPRRLARSARKR LARRKARLNHLKHLIANEFKLNYEDYQSFDESLAKAYKGSLISPYELRFRALNELL SKQDFARVILHIAKRRGYDDIKNSDDKEKGAILKAIKQNEEKLANYQSVGEYLYKE YFQKFKENSKEFTNVRNKKESYERCIAQSFLKDELKLIFKKQREFGFSFSKKFEEEV LSVAFYKRALKDFSHLVGNCSFFTDEKRAPKNSPLAFMFVALTRIINLLNNLKNTE GILYTKDDLNALLNEVLKNGTLTYKQTKKLLGLSDDYEFKGEKGTYFIEFKKYKEF IKALGEHNLSQDDLNEIAKDITLIKDEIKLKKALAKYDLNQNQIDSLSKLEFKDHLN ISFKALKLVTPLMLEGKKYDEACNELNLKVAINEDKKDFLPAFNETYYKDEVTNP VVLRAIKEYRKVLNALLKKYGKVHKINIELAREVGKNHSQRAKIEKEQNENYKAK KDAELECEKLGLKINSKNILKLRLFKEQKEFCAYSGEKIKISDLQDEKMLEIDHIYP YSRSFDDSYMNKVLVFTKQNQEKLNQTPFEAFGNDSAKWQKIEVLAKNLPTKKQ KRILDKNYKDKEQKNFKDRNLNDTRYIARLVLNYTKDYLDFLPLSDDENTKLNDT QKGSKVHVEAKSGMLTSALRHTWGFSAKDRNNHLHHAIDAVIIAYANNSIVKAFS DFKKEQESNSAELYAKKISELDYKNKRKFFEPFSGFRQKVLDKIDEIFVSKPERKKP SGALHEETFRKEEEFYQSYGGKEGVLKALELGKIRKVNGKIVKNGDMFRVDIFKH KKTNKFYAVPIYTMDFALKVLPNKAVARSKKGEIKDWILMDENYEFCFSLYKDSLI LIQTKDMQEPEFVYYNAFTSSTVSLIVSKHDNKFETLSKNQKILFKNANEKEVIAKS IGIQNLKVFEKYIVSALGEVTKAEFRQREDFKK	41
GeoCas9 *G. stearothermophilus* 1087 AA 127 kDa	MRYKIGLDIGITSVGWAVMNLDIPRIEDLGVRIFDRAENPQTGESLALPRRLARSAR RRLRRRKHRLERIRRLVIREGILTKEELDKLFEEKHEIDVWQLRVEALDRKLNNDE LARVLLHLAKRRGFKSNRKSERSNKENSTMLKHIEENRAILSSYRTVGEMIVKDPK FALHKRNKGENYTNTIARDDLEREIRLIFSKQREFGNMSCTEEFENEYITIWASQRP VASKDDIEKKVGFCTFEPKEKRAPKATYTFQSFIAWEHINKLRLISPSGARGLTDEE RRLLYEQAFQKNKITYHDIRTLLHLPDDTYFKGIVYDRGESRKQNENIRFLELDAY HQIRKAVDKVYGKGKSSSFLPIDFDTFGYALTLFKDDADIHSYLRNEYEQNGKRMP NLANKVYDNELIEELLNLSFTKFGHLSLKALRSILPYMEQGEVYSSACERAGYTFT GPKKKQKTMLLPNIPPIANPVVMRALTQARKVVNAIIKKYGSPVSIHIELARDLSQT FDERRKTKKEQDENRKKNETAIRQLMEYGLTLNPTGHDIVKFKLWSEQNGRCAYS LQPIEIERLLEPGYVEVDHVIPYSRSLDDSYTNKVLVLTRENREKGNRIPAEYLGVG TERWQQFETFVLTNKQFSKKKRDRLLRLHYDENEETEFKNRNLNDTRYISRFFANF IREHLKFAESDDKQKVYTVNGRVTAHLRSRWEFNKNREESDLHHAVDAVIVACTT PSDIAKVTAFYQRREQNKELAKKTEPHFPQPWPHFADELRARLSKHPKESIKALNL GNYDDQKLESLQPVFVSRMPKRSVTGAAHQETLRRYVGIDERSGKIQTVVKTKLS EIKLDASGHFPMYGKESDPRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGEPGPV IRTVKIIDTKNQVIPLNDGKTVAYNSNIVRVDVFEKDGKYYCVPVYTMDIMKGILP NKAIEPNKPYSEWKEMTEDYTFRFSLYPNDLIRIELPREKTVKTAAGEEINVKDVFV YYKTIDSANGGLELISHDHRFSLRGVGSRTLKRFEKYQVDVLGNIYKVRGEKRVGL ASSAHSKPGKTIRPLQSTRD	42
LbaCas12a	MSKLEKFTNCYSLSKTLRFKAIPVGKTQENIDNKRLLVEDEKRAEDYKGVKKLLD RYYLSFINDVLHSIKLKNLNNYISLFRKKTRTEKENKELENLEINLRKEIAKAFKGN	43

Description	Sequence	SEQ ID NO:
L. bacterium 1228 AA 143.9 kDa	EGYKSLFKKDIIETILPEFLDDKDEIALVNSFNGFTTAFTGFFDNRENMFSEEAKSTSI AFRCINENLTRYISNMDIFEKVDAIFDKHEVQEIKEKILNSDYDVEDFFEGEFFNFVL TQEGIDVYNAIIGGFVTESGEKIKGLNEYINLYNQKTKQKLPKFKPLYKQVLSDRES LSFYGEGYTSDEEVLEVFRNTLNKNSEIFSSIKKLEKLFKNFDEYSSAGIFVKNGPAI STISKDIFGEWNVIRDKWNAEYDDIHLKKKAVVTEKYEDDRRKSFKKIGSFSLEQL QEYADADLSVVEKLKEIIIQKVDEIYKVYGSSEKLFDADFVLEKSLKKNDAVVAIM KDLLDSVKSFENYIKAFFGEGKETNRDESFYGDFVLAYDILLKVDHIYDAIRNYVT QKPYSKDKFKLYFQNPQFMGGWDKDKETDYRATILRYGSKYYLAIMDKKYAKCL QKIDKDDVNGNYEKINYKLLPGPNKMLPKVFFSKKWMAYYNPSEDIQKIYKNGTF KKGDMFNLNDCHKLIDFFKDSISRYPKWSNAYDFNFSETEKYKDIAGFYREVEEQ GYKVSFESASKKEVDKLVEEGKLYMFQIYNKDFSDKSHGTPNLHTMYFKLLFDEN NHGQIRLSGGAELFMRRASLKKEELVVHPANSPIANKNPDNPKKTTTLSYDVYKD KRFSEDQYELHIPIAINKCPKNIFKINTEVRVLLKHDDNPYVIGIDRGERNLLYIVVV DGKGNIVEQYSLNEIINNFNGIRIKTDYHSLLDKKEKERFEARQNWTSIENIKELKA GYISQVVHKICELVEKYDAVIALEDLNSGFKNSRVKVEKQVYQKFEKMLIDKLNY MVDKKSNPCATGGALKGYQITNKFESFKSMSTQNGFIFYIPAWLTSKIDPSTGFVN LLKTKYTSIADSKKFISSFDRIMYVPEEDLFEFALDYKNFSRTDADYIKKWKLYSYG NRIRIFRNPKKNNVFDWEEVCLTSAYKELFNKYGINYQQGDIRALLCEQSDKAFYS SFMALMSLMLQMRNSITGRTDVDFLISPVKNSDGIFYDSRNYEAQENAILPKNADA NGAYNIARKVLWAIGQFKKAEDEKLDKVKIAISNKEWLEYAQTSVKH	
BhCas12b *B. hisashii* 1108 AA 130.4 kDa	MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQEAIYEHHEQDPKN PKKVSKAEIQAELWDFVLKMQKCNSFTHEVDKDEVFNILRELYEELVPSSVEKKG EANQLSNKFLYPLVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKKKWEED KKKDPLAKILGKLAEYGLIPLFIPYTDSNEPIVKEIKWMEKSRNQSVRRLDKDMFIQ ALERFLSWESWNLKVKEEYEKVEKEYKTLEERIKEDIQALKALEQYEKERQEQLL RDTLNTNEYRLSKRGLRGWREIIQKWLKMDENEPSEKYLEVFKDYQRKHPREAG DYSVYEFLSKKENHFIWRNHPEYPYLYATFCEIDKKKKDAKQQATFTLADPINHPL WVRFEERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPTESGGWEEKGKVDIV LLPSRQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFDRDHLRRYPHK VESGNVGRIYFNMTVNIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIKDSKGK KLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHR ASFNIKLPGETLVKSREVLRKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRV TKWISRQENSDVPLVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEIGKEVKH WRKSLSDGRKGLYGISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQRFAIDQL NHLNALKEDRLKKMANTIIMHALGYCYDVRKKKWQAKNPACQIILFEDLSNYNP YEERSRFENSKLMKWSRREIPRQVALQGEIYGLQVGEVGAQFSSRFHAKTGSPGIR CSVVTKEKLQDNRFFKNLQREGRLTLDKIAVLKEGDLYPDKGGEKFISLSKDRKCV TTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVDGQTVYIPESKDQKQKIIEEFG EGYFILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSDILKDSFDLASELKGEKLML YRDPSGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQSM	44

G. Cas9 equivalents

[352] In some embodiments, the prime editors utilized in the methods and compositions described herein can include any Cas9 equivalent. As used herein, the term "Cas9 equivalent" is a broad term that encompasses any napDNAbp protein that serves the same function as Cas9 in the prime editors despite that its amino acid primary sequence and/or its three-dimensional structure may be different and/or unrelated from an evolutionary standpoint. Thus, while Cas9 equivalents include any Cas9 ortholog, homolog, mutant, or variant described or embraced

herein that are evolutionarily related, the Cas9 equivalents also embrace proteins that may have evolved through convergent evolution processes to have the same or similar function as Cas9, but that do not necessarily have any similarity with regard to amino acid sequence and/or three-dimensional structure. The prime editors utilized in the methods and compositions described here embrace any Cas9 equivalent that would provide the same or similar function as Cas9 despite that the Cas9 equivalent may be based on a protein that arose through convergent evolution. For instance, if Cas9 refers to a type II enzyme of the CRISPR-Cas system, a Cas9 equivalent can refer to a type V or type VI enzyme of the CRISPR-Cas system.

[353] For example, Cas12e (CasX) is a Cas9 equivalent that reportedly has the same function as Cas9 but which evolved through convergent evolution. Thus, the Cas12e (CasX) protein described in Liu *et al.*, "CasX enzymes comprises a distinct family of RNA-guided genome editors," *Nature*, 2019, Vol.566: 218-223, is contemplated to be used with the prime editors utilized in the methods and compositions described herein. In addition, any variant or modification of Cas12e (CasX) is conceivable and within the scope of the present disclosure.

[354] Cas9 is a bacterial enzyme that evolved in a wide variety of species. However, the Cas9 equivalents contemplated herein may also be obtained from archaea, which constitute a domain and kingdom of single-celled prokaryotic microbes different from bacteria.

In some embodiments, Cas9 equivalents may refer to Cas12e (CasX) or Cas12d (CasY), which have been described in, for example, Burstein *et al.*, "New CRISPR–Cas systems from uncultivated microbes." *Cell Res.* 2017 Feb 21. doi: 10.1038/cr.2017.21, the entire contents of which is hereby incorporated by reference. Using genome-resolved metagenomics, a number of CRISPR–Cas systems were identified, including the first reported Cas9 in the archaeal domain of life. This divergent Cas9 protein was found in little-studied nanoarchaea as part of an active CRISPR–Cas system. In bacteria, two previously unknown systems were discovered, CRISPR–Cas12e and CRISPR– Cas12d, which are among the most compact systems yet discovered. In some embodiments, Cas9 refers to Cas12e, or a variant of Cas12e. In some embodiments, Cas9 refers to a Cas12d, or a variant of Cas12d. It should be appreciated that other RNA-guided DNA binding proteins may be used as a nucleic acid programmable DNA binding protein (napDNAbp) and are within the scope of this disclosure. Also see Liu *et al.*, "CasX enzymes comprises a distinct family of RNA-guided genome editors," *Nature*, 2019, Vol.566: 218-223. Any of these Cas9 equivalents are contemplated.

[355] In some embodiments, the Cas9 equivalent comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring Cas12e (CasX) or Cas12d (CasY) protein. In some embodiments, the napDNAbp is a naturally-occurring Cas12e (CasX) or Cas12d (CasY) protein. In some embodiments, the napDNAbp comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a wild-type Cas moiety or any Cas moiety provided herein.

[356] In various embodiments, the nucleic acid programmable DNA binding proteins include, without limitation, Cas9 (*e.g.*, dCas9 and nCas9), Cas12e (CasX), Cas12d (CasY), Cas12a (Cpf1), Cas12b1 (C2c1), Cas13a (C2c2), Cas12c (C2c3), Argonaute, and Cas12b1. One example of a nucleic acid programmable DNA-binding protein that has different PAM specificity than Cas9 is Clustered Regularly Interspaced Short Palindromic Repeats from Prevotella and Francisella 1 (i.e., Cas12a (Cpf1)). Similar to Cas9, Cas12a (Cpf1) is also a Class 2 CRISPR effector, but it is a member of type V subgroup of enzymes, rather than the type II subgroup. It has been shown that Cas12a (Cpf1) mediates robust DNA interference with features distinct from Cas9. Cas12a (Cpf1) is a single RNA-guided endonuclease lacking tracrRNA, and it utilizes a T-rich protospacer-adjacent motif (TTN, TTTN, or YTN). Moreover, Cpf1 cleaves DNA via a staggered DNA double-stranded break. Out of 16 Cpf1-family proteins, two enzymes from *Acidaminococcus* and *Lachnospiraceae* are shown to have efficient genome-editing activity in human cells. Cpf1 proteins are known in the art and have been described previously, for example Yamano *et al.*, "Crystal structure of Cpf1 in complex with guide RNA and target DNA." *Cell* (165) 2016, p. 949-962; the entire contents of which is hereby incorporated by reference.

[357] In still other embodiments, the Cas protein may include any CRISPR associated protein, including but not limited to, Cas12a, Cas12b1, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas6, Cas7, Cas8, Cas9 (also known as Csn1 and Csx12), Cas10, Csy1, Csy2, Csy3, Cse1, Cse2, Csc1, Csc2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, homologs thereof, or modified versions thereof, and preferably comprising a nickase mutation

(e.g., a mutation corresponding to the D10A mutation of the wild type Cas9 polypeptide of SEQ ID NO: 2).

[358] In various other embodiments, the napDNAbp can be any of the following proteins: a Cas9, a Cas12a (Cpf1), a Cas12e (CasX), a Cas12d (CasY), a Cas12b1 (C2c1), a Cas13a (C2c2), a Cas12c (C2c3), a GeoCas9, a CjCas9, a Cas12g, a Cas12h, a Cas12i, a Cas13b, a Cas13c, a Cas13d, a Cas14, a Csn2, an xCas9, an SpCas9-NG, a circularly permuted Cas9, or an Argonaute (Ago) domain, or a variant thereof.

[359] Exemplary Cas9 equivalent protein sequences can include the following:

Description	Sequence
AsCas12a (previously known as Cpf1)*Acidaminococcus sp. (strain BV3L6)*UniProtKB U2UMQ6	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPIIDRIYKTYAD QCLQLVQLDWENLSAAIDSYRKEKTEETRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRH AEIYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVFSAEDIST AIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQ TQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLS FILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSIDLTHIFISHKKLETISSALCDH WDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILS HAHAALDQPLPTTLKKQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPEFSARLTGIKLE MEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTLASGWDVNKEKNNGAILFVKNGLYYL GIMPKQKGRYKALSFEPTEKTSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPI LLSNNFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSKYT KTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFA KGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKK LKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVSHEIIKDRRFTSDKFFFHV PITLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGERNLIYITVIDSTGKILEQRSLNTIQQF DYQKKLDNREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVVLENLNFG FKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTSFAKMGTQ SGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMN RNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIENHRFTGRYRDLYPANEL IALLEEKGIVFRDGSNILPKLLENDDSHAIDTMVALIRSVLQMRNSNAATGEDYINSPVRDL NGVCFDSRFQNPEWPMDADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQE LRN (SEQ ID NO: 45)
AsCas12a nickase (e.g., R1226A)	MTQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDHYKELKPIIDRIYKTYAD QCLQLVQLDWENLSAAIDSYRKEKTEETRNALIEEQATYRNAIHDYFIGRTDNLTDAINKRH AEIYKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFDKFTTYFSGFYENRKNVFSAEDIST AIPHRIVQDNFPKFKENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSIEEVFSFPFYNQLLTQ TQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLS FILEEFKSDEEVIQSFCKYKTLLRNENVLETAEALFNELNSIDLTHIFISHKKLETISSALCDH WDTLRNALYERRISELTGKITKSAKEKVQRSLKHEDINLQEIISAAGKELSEAFKQKTSEILS HAHAALDQPLPTTLKKQEEKEILKSQLDSLLGLYHLLDWFAVDESNEVDPEFSARLTGIKLE MEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTLASGWDVNKEKNNGAILFVKNGLYYL GIMPKQKGRYKALSFEPTEKTSEGFDKMYYDYFPDAAKMIPKCSTQLKAVTAHFQTHTTPI LLSNNFIEPLEITKEIYDLNNPEKEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSKYT KTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKEIMDAVETGKLYLFQIYNKDFA KGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAELFYRPKSRMKRMAHRLGEKMLNKK LKDQKTPIPDTLYQELYDYVNHRLSHDLSDEARALLPNVITKEVSHEIIKDRRFTSDKFFFHV PITLNYQAANSPSKFNQRVNAYLKEHPETPIIGIDRGERNLIYITVIDSTGKILEQRSLNTIQQF DYQKKLDNREKERVAARQAWSVVGTIKDLKQGYLSQVIHEIVDLMIHYQAVVVLENLNFG FKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGVLNPYQLTDQFTSFAKMGTQ

	SGFLFYVPAPYTSKIDPLTGFVDPFVWKTIKNHESRKHFLEGFDFLHYDVKTGDFILHFKMN RNLSFQRGLPGFMPAWDIVFEKNETQFDAKGTPFIAGKRIVPVIENHRFTGRYRDLYPANEL IALLEEKGIVFRDGSNILPKLLENDDSHAIDTMVALIRSVLQMANSNAATGEDYINSPVRDL NGVCFDSRFQNPEWPMDADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQDWLAYIQE LRN (SEQ ID NO: 46)
LbCas12a (previously known as Cpf1)*Lachnospiraceae bacterium GAM79* Ref Seq. WP_119623382.1	MNYKTGLEDFIGKESLSKTLRNALIPTESTKIHMEEMGVIRDDELRAEKQQELKEIMDDYYR TFIEEKLGQIQGIQWNSLFQKMEETMEDISVRKDLDKIQNEKRKEICCYFTSDKRFKDLFNA KLITDILPNFIKDNKEYTEEEKAEKEQTRVLFQRFATAFTNYFNQRRNNFSEDNISTAISFRIV NENSEIHLQNMRAFQRIEQQYPEEVCGMEEEYKDMLQEWQMKHIYSVDFYDRELTQPGIE YYNGICGKINEHMNQFCQKNRINKNDFRMKKLHKQILCKKSSYYEIPFRFESDQEVYDALN EFIKTMKKKEIIRRCVHLGQECDDYDLGKIYISSNKYEQISNALYGSWDTIRKCIKEEYMDA LPGKGEKKEEKAEAAAKKEEYRSIADIDKIISLYGSEMDRTISAKKCITEICDMAGQISIDPLV CNSDIKLLQNKEKTTEIKTILDSFLHVYQWGQTFIVSDIIEKDSYFYSELEDVLEDFEGITTLY NHVRSYVTQKPYSTVKFKLHFGSPTLANGWSQSKEYDNNAILLMRDQKFYLGIFNVRNKP DKQIIKGHEKEEKGDYKKMIYNLLPGPSKMLPKVFITSRSGQETYKPSKHILDGYNEKRHIK SSPKFDLGYCWDLIDYYKECIHKHPDWKNYDFHFSDTKDYEDISGFYREVEMQGYQIKWT YISADEIQKLDEKGQIFLFQIYNKDFSVHSTGKDNLHTMYLKNLFSEENLKDIVLKLNGEAE LFFRKASIKTPIVHKKGSVLVNRSYTQTVGNKEIRVSIPEEYYTEIYNYLNHIGKGKLSSEAQ RYLDEGKIKSFTATKDIVKNYRYCCDHYFLHLPITINFKAKSDVAVNERTLAYIAKKEDIHII GIDRGERNLLYISVVDVHGNIREQRSFNIVNGYDYQQKLKDREKSRDAARKNWEEIEKIKE LKEGYLSMVIHYIAQLVVKYNAVVAMEDLNYGFKTGRFKVERQVYQKFETMLIEKLHYLV FKDREVCEEGGVLRGYQLTYIPESLKKVGKQCGFIFYVPAGYTSKIDPTTGFVNLFSFKNLT NRESRQDFVGKFDEIRYDRDKKMFEFSFDYNNYIKKGTILASTKWKVYTNGTRLKRIVVNG KYTSQSMEVELTDAMEKMLQRAGIEYHDGKDLKGQIVEKGIEAEIIDIFRLTVQMRNSRSES EDREYDRLISPVLNDKGEFFDTATADKTLPQDADANGAYCIALKGLYEVKQIKENWKENE QFPRNKLVQDNKTWFDFMQKKRYL (SEQ ID NO: 47)
PcCas12a – previously known at Cpf1*Prevotella copri* Ref Seq. WP_119227726.1	MAKNFEDFKRLYSLSKTLRFEAKPIGATLDNIVKSGLLDEDEHRAASYVKVKKLIDEYHKV FIDRVLDDGCLPLENKGNNNSLAEYYESYVSRAQDEDAKKKFKEIQQNLRSVIAKKLTEDK AYANLFGNKLIESYKDKEDKKKIIDSDLIQFINTAESTQLDSMSQDEAKELVKEFWGFVTYF YGFFDNRKNMYTAEEKSTGIAYRLVNENLPKFIDNIEAFNRAITRPEIQENMGVLYSDFSEY LNVESIQEMFQLDYYNMLLTQKQIDVYNAIIGGKTDDEHDVKIKGINEYINLYNQQHKDDK LPKLKALFKQILSDRNAISWLPEEFNSDQEVLNAIKDCYERLAENVLGDKVLKSLLGSLADY SLDGIFIRNDLQLTDISQKMFGNWGVIQNAIMQNIKRVAPARKHKESEEDYEKRIAGIFKKA DSFSISYINDCLNEADPNNAYFVENYFATFGAVNTPTMQRENLFALVQNAYTEVAALLHSD YPTVKHLAQDKANVSKIKALLDAIKSLQHFVKPLLGKGDESDKDERFYGELASLWAELDT VTPLYNMIRNYMTRKPYSQKKIKLNFENPQLLGGWDANKEKDYATIILRRNGLYYLAIMD KDSRKLLGKAMPSDGECYEKMVYKFFKDVTTMIPKCSTQLKDVQAYFKVNTDDYVLNSK AFNKPLTITKEVFDLNNVLYGKYKKFQKGYLTATGDNVGYTHAVNVWIKFCMDFLNSYDS TCIYDFSSLKPESYLSLDAFYQDANLLLYKLSFARASVSYINQLVEEGKMYLFQIYNKDFSE YSKGTPNMHTLYWKALFDERNLADVVYKLNGQAEMFYRKKSIENTHPTHPANHPILNKNK DNKKKESLFDYDLIKDRRYTVDKFMFHVPITMNFKSVGSENINQDVKAYLRHADDMHIIGI DRGERHLLYLVVIDLQGNIKEQYSLNEIVNEYNGNTYHTNYHDLLDVREEERLKARQSWQ TIENIKELKEGYLSQVIHKITQLMVRYHAIVVLEDLSKGFMRSRQKVEKQVYQKFEKMLID KLNYLVDKKTDVSTPGGLLNAYQLTCKSDSSQKLGKQSGFLFYIPAWNTSKIDPVTGFVNL LDTHSLNSKEKIKAFFSKFDAIRYNKDKKWFEFNLDYDKFGKKAEDTRTKWTLCTRGMRI DTFRNKEKNSQWDNQEVDLTTEMKSLLEHYYIDIHGNLKDAISAQTDKAFFTGLLHILKLT LQMRNSITGTETDYLVSPVADENGIFYDSRSCGNQLPENADANGAYNIARKGLMLIEQIKN AEDLNNVKFDISNKAWLNFAQQKPYKNG (SEQ ID NO: 48)
ErCas12a – previously known at Cpf1*Eubacterium rectale*	MFSAKLISDILPEFVIHNNNYSASEKEEKTQVIKLFSRFATSFKDYFKNRANCFSANDISSSSC HRIVNDNAEIFFSNALVYRRIVKNLSNDDINKISGDMKDSLKEMSLEEIYSYEKYGEFITQEG ISFYNDICGKVNLFMNLYCQKNKENKNLYKLRKLHKQILCIADTSYEVPYKFESDEEVYQS VNGFLDNISSKHIVERLRKIGENYNGYNLDKIYIVSKFYESVSQKTYRDWETINTALEIHYNN ILPGNGKSKADKVKKAVKNDLQKSITEINELVSNYKLCPDDNIKAETYIHEISHILNNFEAQE LKYNPEIHLVESELKASELKNVLDVIMNAFHWCSVFMTEELVDKDNNFYAELEEIYDEIYP VISLYNLVRNYVTQKPYSTKKIKLNFGIPTLADGWSKSKEYSNNAIILMRDNLYYLGIFNAK

Ref Seq. WP_119223642.1	NKPDKKIIEGNTSENKGDYKKMIYNLLPGPNKMIPKVFLSSKTGVETYKPSAYILEGYKQNK HLKSSKDFDITFCHDLIDYFKNCIAIHPEWKNFGFDFSDTSTYEDISGFYREVELQGYKIDWT YISEKDIDLLQEKGQLYLFQIYNKDFSKKSSGNDNLHTMYLKNLFSEENLKDIVLKLNGEAE IFFRKSSIKNPIIHKKGSILVNRTYEAEEKDQFGNIQIVRKTIPENIYQELYKYFNDKSDKELSD EAAKLKNVVGHHEAATNIVKDYRYTYDKYFLHMPITINFKANKTSFINDRILQYIAKEKDL HVIGIDRGERNLIYVSVIDTCGNIVEQKSFNIVNGYDYQIKLKQQEGARQIARKEWKEIGKIK EIKEGYLSLVIHEISKMVIKYNAIIAMEDLSYGFKKGRFKVERQVYQKFETMLINKLNYLVF KDISITENGGLLKGYQLTYIPDKLKNVGHQCGCIFYVPAAYTSKIDPTTGFVNIFKFKDLTVD AKREFIKKFDSIRYDSDKNLFCFTFDYNNFITQNTVMSKSSWSVYTYGVRIKRRFVNGRFSN ESDTIDITKDMEKTLEMTDINWRDGHDLRQDIIDYEIVQHIFEIFKLTVQMRNSLSELEDRDY DRLISPVLNENNIFYDSAKAGDALPKDADANGAYCIALKGLYEIKQITENWKEDGKFSRDK LKISNKDWFDFIQNKRYL (SEQ ID NO: 49)
CsCas12a – previously known at Cpf1*Clostridium sp. AF34-10BH*Ref Seq. WP_118538418.1	MNYKTGLEDFIGKESLSKTLRNALIPTESTKIHMEEMGVIRDDELRAEKQQELKEIMDDYYR AFIEEKLGQIQGIQWNSLFQKMEETMEDISVRKDLDKIQNEKRKEICCYFTSDKRFKDLFNA KLITDILPNFIKDNKEYTEEEKAEKEQTRVLFQRFATAFTNYFNQRRNNFSEDNISTAISFRIV NENSEIHLQNMRAFQRIEQQQYPEEVCGMEEEYKDMLQEWQMKHIYLVDFYDRVLTQPGIE YYNGICGKINEHMNQFCQKNRINKNDFRMKKLHKQILCKKSSYYEIPFRFESDQEVYDALN EFIKTMKEKEIICRCVHLGQKCDDYDLGKIYISSNKYEQISNALYGSWDTIRKCIKEEYMDA LPGKGEKKEEKAEAAAKKEEYRSIADIDKIISLYGSEMDRTISAKKCITEICDMAGQISTDPL VCNSDIKLLQNKEKTTEIKTILDSFLHVYQWGQTFIVSDIIEKDSYFYSELEDVLEDFEGITTL YNHVRSYVTQKPYSTVKFKLHFGSPTLANGWSQSKEYDNNAILLMRDQKFYLGIFNVRNK PDKQIIKGHEKEEKGDYKKMIYNLLPGPSKMLPKVFITSRSGQETYKPSKHILDGYNEKRHI KSSPKFDLGYCWDLIDYYKECIHKHPDWKNYDFHFSDTKDYEDISGFYREVEMQGYQIKW TYISADEIQKLDEKGQIFLFQIYNKDFSVHSTGKDNLHTMYLKNLFSEENLKDIVLKLNGEA ELFFRKASIKTPVVHKKGSVLVNRSYTQTVGDKEIRVSIPEEYYTEIYNYLNHIGRGKLSTEA QRYLEERKIKSFTATKDIVKNYRYCCDHYFLHLPITINFKAKSDIAVNERTLAYIAKKEDIHII GIDRGERNLLYISVVDVHGNIREQRSFNIVNGYDYQQKLKDREKSRDAARKNWEEIEKIKE LKEGYLSMVIHYIAQLVVKYNAVVAMEDLNYGFKTGRFKVERQVYQKFETMLIEKLNYLV FKDREVCEEGGVLRGYQLTYIPESLKKVGKQCGFIFYVPAGYTSKIDPTTGFVNLFSFKNLT NRESRQDFVGKFDEIRYDRDKKMFEFSFDYNNYIKKGTMLASTKWKVYTNGTRLKRIVVN GKYTSQSMEVELTDAMEKMLQRAGIEYHDGKDLKGQIVEKGIEAEIIDIFRLTVQMRNSRS ESEDREYDRLISPVLNDKGEFFDTATADKTLPQDADANGAYCIALKGLYEVKQIKENWKEN EQFPRNKLVQDNKTWFDFMQKKRYL (SEQ ID NO: 50)
BhCas12b *Bacillus hisashii* Ref Seq. WP_095142515.1	MATRSFILKIEPNEEVKKGLWKTHEVLNHGIAYYMNILKLIRQEAIYEHHEQDPKNPKKVSK AEIQAELWDFVLKMQKCNSFTHEVDKDEVFNILRELYEELVPSSVEKKGEANQLSNKFLYP LVDPNSQSGKGTASSGRKPRWYNLKIAGDPSWEEEKKKWEEDKKKDPLAKILGKLAEYGL IPLFIPYTDSNEPIVKEIKWMEKSRNQSVRRLDKDMFIQALERFLSWESWNLKVKEEYEKVE KEYKTLEERIKEDIQALKALEQYEKERQEQLLRDTLNTNEYRLSKRGLRGWREIIQKWLKM DENEPSEKYLEVFKDYQRKHPREAGDYSVYEFLSKKENHFIWRNHPEYPYLYATFCEIDKK KKDAKQQATFTLADPINHPLWVRFEERSGSNLNKYRILTEQLHTEKLKKKLTVQLDRLIYPT ESGGWEEKGKVDIVLLPSRQFYNQIFLDIEEKGKHAFTYKDESIKFPLKGTLGGARVQFDRD HLRRYPHKVESGNVGRIYFNMTVNIEPTESPVSKSLKIHRDDFPKVVNFKPKELTEWIKDSK GKKLKSGIESLEIGLRVMSIDLGQRQAAAASIFEVVDQKPDIEGKLFFPIKGTELYAVHRASF NIKLPGETLVKSREVLRKAREDNLKLMNQKLNFLRNVLHFQQFEDITEREKRVTKWISRQE NSDVPLVYQDELIQIRELMYKPYKDWVAFLKQLHKRLEVEIGKEVKHWRKSLSDGRKGLY GISLKNIDEIDRTRKFLLRWSLRPTEPGEVRRLEPGQRFAIDQLNHLNALKEDRLKKMANTII MHALGYCYDVRKKKWQAKNPACQIILFEDLSNYNPYEERSRFENSKLMKWSRREIPRQVA LQGEIYGLQVGEVGAQFSSRFHAKTGSPGIRCSVVTKEKLQDNRFFKNLQREGRLTLDKIAV LKEGDLYPDKGGEKFISLSKDRKCVTTHADINAAQNLQKRFWTRTHGFYKVYCKAYQVD GQTVYIPESKDQKQKIIEEFGEGYFILKDGVYEWVNAGKLKIKKGSSKQSSSELVDSDILKDS FDLASELKGEKLMLYRDPSGNVFPSDKWMAAGVFFGKLERILISKLTNQYSISTIEDDSSKQ SM (SEQ ID NO: 44)
ThCas12b *Thermomonas hydrothermalis*	MSEKTTQRAYTLRLNRASGECAVCQNNSCDCWHDALWATHKAVNRGAKAFGDWLLTLR GGLCHTLVEMEVPAKGNNPPQRPTDQERRDRRVLLALSWLSVEDEHGAPKEFIVATGRDS ADDRAKKVEEKLREILEKRDFQEHEIDAWLQDCGPSLKAHIREDAVWVNRRALFDAAVER

Ref Seq. WP_072754838	IKTLTWEEAWDFLEPFFGTQYFAGIGDGKDKDDAEGPARQGEKAKDLVQKAGQWLSARF GIGTGADFMSMAEAYEKIAKWASQAQNGDNGKATIEKLACALRPSEPPTLDTVLKCISGPG HKSATREYLKTLDKKSTVTQEDLNQLRKLADEDARNCRKKVGKKGKKPWADEVLKDVE NSCELTYLQDNSPARHREFSVMLDHAARRVSMAHSWIKKAEQRRRQFESDAQKLKNLQER APSAVEWLDRFCESRSMTTGANTGSGYRIRKRAIEGWSYVVQAWAEASCDTEDKRIAAAR KVQADPEIEKFGDIQLFEALAADEAICVWRDQEGTQNPSILIDYVTGKTAEHNQKRFKVPA YRHPDELRHPVFCDFGNSRWSIQFAIHKEIRDRDKGAKQDTRQLQNRHGLKMRLWNGRSM TDVNLHWSSKRLTADLALDQNPNPNPTEVTRADRLGRAASSAFDHVKIKNVFNEKEWNGR LQAPRAELDRIAKLEEQGKTEQAEKLRKRLRWYVSFSPCLSPSGPFIVYAGQHNIQPKRSGQ YAPHAQANKGRARLAQLILSRLPDLRILSVDLGHRFAAACAVWETLSSDAFRREIQGLNVL AGGSGEGDLFLHVEMTGDDGKRRTVVYRRIGPDQLLDNTPHPAPWARLDRQFLIKLQGED EGVREASNEELWTVHKLEVEVGRTVPLIDRMVRSGFGKTEKQKERLKKLRELGWISAMPN EPSAETDEKEGEIRSISRSVDELMSSALGTLRLALKRHGNRARIAFAMTADYKPMPGGQKY YFHEAKEASKNDDETKRRDNQIEFLQDALSLWHDLFSSPDWEDNEAKKLWQNHIATLPNY QTPEEISAELKRVERNKKRKENRDKLRTAAKALAENDQLRQHLHDTWKERWESDDQQWK ERLRSLKDWIFPRGKAEDNPSIRHVGGLSITRINTISGLYQILKAFKMRPEPDDLRKNIPQKG DDELENFNRRLLEARDRLREQRVKQLASRIIEAALGVGRIKIPKNGKLPKRPRTTVDTPCHA VVIESLKTYRPDDLRTRRENRQLMQWSSAKVRKYLKEGCELYGLHFLEVPANYTSRQCSR TGLPGIRCDDVPTGDFLKAPWWRRAINTAREKNGGDAKDRFLVDLYDHLNNLQSKGEALP ATVRVPRQGGNLFIAGAQLDDTNKERRAIQADLNAAANIGLRALLDPDWRGRWWYVPCK DGTSEPALDRIEGSTAFNDVRSLPTGDNSSRRAPREIENLWRDPSGDSLESGTWSPTRAYWD TVQSRVIELLRRHAGLPTS (SEQ ID NO: 51)
LsCas12b*Lacevella sacchari*WP_132221894.1	MSIRSFKLKLKTKSGVNAEQLRRGLWRTHQLINDGIAYYMNWLVLLRQEDLFIRNKETNEI EKRSKEEIQAVLLERVHKQQQRNQWSGEVDEQTLLQALRQLYEEIVPSVIGKSGNASLKAR FFLGPLVDPNNKTTKDVSKSGPTPKWKKMKDAGDPNWVQEYEKYMAERQTLVRLEEMG LIPLFPMYTDEVGDIHWLPQASGYTRTWDRDMFQQAIERLLSWESWNRRVRERRAQFEKK THDFASRFSESDVQWMNKLREYEAQQEKSLEENAFAPNEPYALTKKALRGWERVYHSWM RLDSAASEEAYWQEVATCQTAMRGEFGDPAIYQFLAQKENHDIWRGYPERVIDFAELNHL QRELRRAKEDATFTLPDSVDHPLWVRYEAPGGTNIHGYDLVQDTKRNLTLILDKFILPDEN GSWHEVKKVPFSLAKSKQFHRQVWLQEEQKQKKREVVFYDYSTNLPHLGTLAGAKLQWD RNFLNKRTQQQIEETGEIGKVFFNISVDVRPAVEVKNGRLQNGLGKALTVLTHPDGTKIVTG WKAEQLEKWVGESGRVSSLGLDSLSEGLRVMSIDLGQRTSATVSVFEITKEAPDNPYKFFY QLEGTEMFAVHQRSFLLALPGENPPQKIKQMREIRWKERNRIKQQVDQLSAILRLHKKVNE DERIQAIDKLLQKVASWQLNEEIATAWNQALSQLYSKAKENDLQWNQAIKNAHHQLEPVV GKQISLWRKDLSTGRQGIAGLSLWSIEELEATKKLLTRWSKRSREPGVVKRIERFETFAKQI QHHINQVKENRLKQLANLIVMTALGYKYDQEQKKWIEVYPACQVVLFENLRSYRFSFERS RRENKKLMEWSHRSIPKLVQMQGELFGLQVADVYAAYSSRYHGRTGAPGIRCHALTEADL RNETNIIHELIEAGFIKEEHRPYLQQGDLVPWSGGELFATLQKPYDNPRILTLHADINAAQNI QKRFWHPSMWFRVNCESVMEGEIVTYVPKNKTVHKKQGKTFRFVKVEGSDVYEWAKWS KNRNKNTFSSITERKPPSSMILFRDPSGTFFKEQEWVEQKTFWGKVQSMIQAYMKKTIVQR MEE (SEQ ID NO: 52)
DtCas12b*Dsulfonatronum thiodismutans* WP_031386437	MVLGRKDDTAELRRALWTTHEHVNLAVAEVERVLLRCRGRSYWTLDRRGDPVHVPESQV AEDALAMAREAQRRNGWPVVGEDEEILLALRYLYEQIVPSCLLDDLGKPLKGDAQKIGTN YAGPLFDSDTCRRDEGKDVACCGPFHEVAGKYLGALPEWATPISKQEFDGKDASHLRFKA TGGDDAFFRVSIEKANAWYEDPANQDALKNKAYNKDDWKKEKDKGISSWAVKYIQKQL QLGQDPRTEVRRKLWLELGLLPLFIPVFDKTMVGNLWNRLAVRLALAHLLSWESWNHRA VQDQALARAKRDELAALFLGMEDGFAGLREYELRRNESIKQHAFEPVDRPYVVSGRALRS WTRVREEWLRHGDTQESRKNICNRLQDRLRGKFGDPDVFHWLAEDGQEALWKERDCVTS FSLLNDADGLLEKRKGYALMTFADARLHPRWAMYEAPGGSNLRTYQIRKTENGLWADVV LLSPRNESAAVEEKTFNVRLAPSGQLSNVSFDQIQKGSKMVGRCRYQSANQQFEGLLGGAE ILFDRKRIANEQHGATDLASKPGHVWFKLTLDVRPQAPQGWLDGKGRPALPPEAKHFKTA LSNKSKFADQVRPGLRVLSVDLGVRSFAACSVFELVRGGPDQGTYFPAADGRTVDDPEKL WAKHERSFKITLPGENPSRKEEIARRAAMEELRSLNGDIRRLKAILRLSVLQEDDPRTEHLRL FMEAIVDDPAKSALNAELFKGFGDDRFRSTPDLWKQHCHFFHDKAEKVVAERFSRWRTET RPKSSSWQDWRERRGYAGGKSYWAVTYLEAVRGLILRWNMRGRTYGEVNRQDKKQFGT

	VASALLHHINQLKEDRIKTGADMIIQAARGFVPRKNGAGWVQVHEPCRLILFEDLARYRFR TDRSRRENSRLMRWSHREIVNEVGMQGELYGLHVDTTEAGFSSRYLASSGAPGVRCRHLV EEDFHDGLPGMHLVGELDWLLPKDKDRTANEARRLLGGMVRPGMLVPWDGGELFATLN AASQLHVIHADINAAQNLQRRFWGRCGEAIRIVCNQLSVDGSTRYEMAKAPKARLLGALQ QLKNGDAPFHLTSIPNSQKPENSYVMTPTNAGKKYRAGPGEKSSGEEDELALDIVEQAEEL AQGRKTFFRDPSGVFFAPDRWLPSEIYWSRIRRRIWQVTLERNSSGRQERAEMDEMPY (SEQ ID NO: 53)

[360] The prime editors utilized in the methods and compositions described herein may also comprise Cas12a (Cpf1) (dCpf1) variants that may be used as a guide nucleotide sequence-programmable DNA-binding protein domain. The Cas12a (Cpf1) protein has a RuvC-like endonuclease domain that is similar to the RuvC domain of Cas9 but does not have a HNH endonuclease domain, and the N-terminal of Cas12a (Cpf1) does not have the alfa-helical recognition lobe of Cas9. It was shown in Zetsche *et al.*, *Cell*, 163, 759–771, 2015 (which is incorporated herein by reference) that, the RuvC-like domain of Cas12a (Cpf1) is responsible for cleaving both DNA strands and inactivation of the RuvC-like domain inactivates Cas12a (Cpf1) nuclease activity.

In some embodiments, the napDNAbp is a single effector of a microbial CRISPR-Cas system. Single effectors of microbial CRISPR-Cas systems include, without limitation, Cas9, Cas12a (Cpf1), Cas12b1 (C2c1), Cas13a (C2c2), and Cas12c (C2c3). Typically, microbial CRISPR-Cas systems are divided into Class 1 and Class 2 systems. Class 1 systems have multi-subunit effector complexes, while Class 2 systems have a single protein effector. For example, Cas9 and Cas12a (Cpf1) are Class 2 effectors. In addition to Cas9 and Cas12a (Cpf1), three distinct Class 2 CRISPR-Cas systems (Cas12b1, Cas13a, and Cas12c) have been described by Shmakov *et al.*, "Discovery and Functional Characterization of Diverse Class 2 CRISPR Cas Systems", *Mol. Cell*, 2015 Nov 5; 60(3): 385–397, the entire contents of which are hereby incorporated by reference.

[361] Effectors of two of the systems, Cas12b1 and Cas12c, contain RuvC-like endonuclease domains related to Cas12a. A third system, Cas13a contains an effector with two predicted HEPN RNase domains. Production of mature CRISPR RNA is tracrRNA-independent, unlike production of CRISPR RNA by Cas12b1. Cas12b1 depends on both CRISPR RNA and tracrRNA for DNA cleavage. Bacterial Cas13a has been shown to possess a unique RNase activity for CRISPR RNA maturation distinct from its RNA-activated single-stranded RNA degradation activity. These RNase functions are different from each other and from the CRISPR RNA-processing behavior of Cas12a. See, e.g., East-Seletsky, *et al.*, "Two distinct RNase

activities of CRISPR-Cas13a enable guide-RNA processing and RNA detection", *Nature,* 2016 Oct 13;538(7624):270-273, the entire contents of which are hereby incorporated by reference. *In vitro* biochemical analysis of Cas13a in *Leptotrichia shahii* has shown that Cas13a is guided by a single CRISPR RNA and can be programed to cleave ssRNA targets carrying complementary protospacers. Catalytic residues in the two conserved HEPN domains mediate cleavage. Mutations in the catalytic residues generate catalytically inactive RNA-binding proteins. See e.g., Abudayyeh *et al.*, "C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector", *Science,* 2016 Aug 5; 353(6299), the entire contents of which are hereby incorporated by reference.

[362] The crystal structure of *Alicyclobaccillus acidoterrastris* Cas12b1 (AacC2c1) has been reported in complex with a chimeric single-molecule guide RNA (sgRNA). See e.g., Liu *et al.*, "C2c1-sgRNA Complex Structure Reveals RNA-Guided DNA Cleavage Mechanism", *Mol. Cell,* 2017 Jan 19;65(2):310-322, the entire contents of which are hereby incorporated by reference. The crystal structure has also been reported in *Alicyclobacillus acidoterrestris* C2c1 bound to target DNAs as ternary complexes. See e.g., Yang *et al.*, "PAM-dependent Target DNA Recognition and Cleavage by C2C1 CRISPR-Cas endonuclease", *Cell,* 2016 Dec 15;167(7):1814-1828, the entire contents of which are hereby incorporated by reference. Catalytically competent conformations of AacC2c1, both with target and non-target DNA strands, have been captured independently positioned within a single RuvC catalytic pocket, with C2c1-mediated cleavage resulting in a staggered seven-nucleotide break of target DNA. Structural comparisons between C2c1 ternary complexes and previously identified Cas9 and Cpf1 counterparts demonstrate the diversity of mechanisms used by CRISPR-Cas9 systems. In some embodiments, the napDNAbp may be a C2c1, a C2c2, or a C2c3 protein. In some embodiments, the napDNAbp is a C2c1 protein. In some embodiments, the napDNAbp is a Cas13a protein. In some embodiments, the napDNAbp is a Cas12c protein. In some embodiments, the napDNAbp comprises an amino acid sequence that is at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring Cas12b1 (C2c1), Cas13a (C2c2), or Cas12c (C2c3) protein. In some embodiments, the napDNAbp is a naturally-occurring Cas12b1 (C2c1), Cas13a (C2c2), or Cas12c (C2c3) protein.

H. Cas9 circular permutants

[363] In various embodiments, the prime editors utilized in the methods and compositions disclosed herein may comprise a circular permutant of Cas9.

[364] The term "circularly permuted Cas9" or "circular permutant" of Cas9 or "CP-Cas9") refers to any Cas9 protein, or variant thereof, that occurs or has been modified or engineered as a circular permutant variant, which means the N-terminus and the C-terminus of a Cas9 protein (e.g., a wild type Cas9 protein) have been topically rearranged. Such circularly permuted Cas9 proteins, or variants thereof, retain the ability to bind DNA when complexed with a guide RNA (gRNA). See, Oakes *et al.*, "Protein Engineering of Cas9 for enhanced function," *Methods Enzymol*, 2014, 546: 491–511 and Oakes *et al.*, "CRISPR-Cas9 Circular Permutants as Programmable Scaffolds for Genome Modification," *Cell*, January 10, 2019, 176: 254-267, each of are incorporated herein by reference. The instant disclosure contemplates any previously known CP-Cas9 or use a new CP-Cas9 so long as the resulting circularly permuted protein retains the ability to bind DNA when complexed with a guide RNA (gRNA).

Any of the Cas9 proteins described herein, including any variant, ortholog, or any engineered or naturally occurring Cas9 or equivalent thereof, may be reconfigured as a circular permutant variant.

In various embodiments, the circular permutants of Cas9 may have the following structure: N-terminus-[original C-terminus] – [optional linker] – [original N-terminus]-C-terminus.

[365] As an example, the present disclosure contemplates the following circular permutants of canonical *S. pyogenes* Cas9 (1368 amino acids of UniProtKB - Q99ZW2 (CAS9_STRP1) (numbering is based on the amino acid position in SEQ ID NO: 2)):

N-terminus-[1268-1368]-[optional linker]-[1-1267]-C-terminus;

N-terminus-[1168-1368]-[optional linker]-[1-1167]-C-terminus;

N-terminus-[1068-1368]-[optional linker]-[1-1067]-C-terminus;

N-terminus-[968-1368]-[optional linker]-[1-967]-C-terminus;

N-terminus-[868-1368]-[optional linker]-[1-867]-C-terminus;

N-terminus-[768-1368]-[optional linker]-[1-767]-C-terminus;

N-terminus-[668-1368]-[optional linker]-[1-667]-C-terminus;

N-terminus-[568-1368]-[optional linker]-[1-567]-C-terminus;

N-terminus-[468-1368]-[optional linker]-[1-467]-C-terminus;

N-terminus-[368-1368]-[optional linker]-[1-367]-C-terminus;
N-terminus-[268-1368]-[optional linker]-[1-267]-C-terminus;
N-terminus-[168-1368]-[optional linker]-[1-167]-C-terminus,
N-terminus-[68-1368]-[optional linker]-[1-67]-C-terminus; or
N-terminus-[10-1368]-[optional linker]-[1-9]-C-terminus, or the corresponding circular permutants of other Cas9 proteins (including other Cas9 orthologs, variants, etc.).

[366] In particular embodiments, the circular permutant Cas9 has the following structure (based on *S. pyogenes* Cas9 (1368 amino acids of UniProtKB - Q99ZW2 (CAS9_STRP1) (numbering is based on the amino acid position in SEQ ID NO: 2):
N-terminus-[102-1368]-[optional linker]-[1-101]-C-terminus;
N-terminus-[1028-1368]-[optional linker]-[1-1027]-C-terminus;
N-terminus-[1041-1368]-[optional linker]-[1-1043]-C-terminus;
N-terminus-[1249-1368]-[optional linker]-[1-1248]-C-terminus; or
N-terminus-[1300-1368]-[optional linker]-[1-1299]-C-terminus, or the corresponding circular permutants of other Cas9 proteins (including other Cas9 orthologs, variants, etc.).

[367] In still other embodiments, the circular permeant Cas9 has the following structure (based on *S. pyogenes* Cas9 (1368 amino acids of UniProtKB - Q99ZW2 (CAS9_STRP1) (numbering is based on the amino acid position in SEQ ID NO: 2):
N-terminus-[103-1368]-[optional linker]-[1-102]-C-terminus;
N-terminus-[1029-1368]-[optional linker]-[1-1028]-C-terminus;
N-terminus-[1042-1368]-[optional linker]-[1-1041]-C-terminus;
N-terminus-[1250-1368]-[optional linker]-[1-1249]-C-terminus; or
N-terminus-[1301-1368]-[optional linker]-[1-1300]-C-terminus, or the corresponding circular permutants of other Cas9 proteins (including other Cas9 orthologs, variants, etc.).

[368] In some embodiments, the circular permutant can be formed by linking a C-terminal fragment of a Cas9 to an N-terminal fragment of a Cas9, either directly or by using a linker, such as an amino acid linker. In some embodiments, The C-terminal fragment may correspond to the C-terminal 95% or more of the amino acids of a Cas9 (e.g., amino acids about 1300-1368), or the C-terminal 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, or 5% or more of a Cas9 (e.g., any one of SEQ ID NOs: 54-63). The N-terminal portion may correspond to the N-terminal 95% or more of the amino acids of a Cas9 (e.g., amino

acids about 1-1300), or the N-terminal 90%, 85%, 80%, 75%, 70%, 65%, 60%, 55%, 50%, 45%, 40%, 35%, 30%, 25%, 20%, 15%, 10%, or 5% or more of a Cas9 (e.g., of SEQ ID NO: 2).

[369] In some embodiments, the circular permutant can be formed by linking a C-terminal fragment of a Cas9 to an N-terminal fragment of a Cas9, either directly or by using a linker, such as an amino acid linker. In some embodiments, the C-terminal fragment that is rearranged to the N-terminus includes or corresponds to the C-terminal 30% or less of the amino acids of a Cas9 (*e.g.*, amino acids 1012-1368 of SEQ ID NO: 2). In some embodiments, the C-terminal fragment that is rearranged to the N-terminus includes or corresponds to the C-terminal 30%, 29%, 28%, 27%, 26%, 25%, 24%, 23%, 22%, 21%, 20%, 19%, 18%, 17%, 16%, 15%, 14%, 13%, 12%, 11%, 10%, 9%, 8%, 7%, 6%, 5%, 4%, 3%, 2%, or 1% of the amino acids of a Cas9 (*e.g.*, the Cas9 of SEQ ID NO: 2). In some embodiments, the C-terminal fragment that is rearranged to the N-terminus includes or corresponds to the C-terminal 410 residues or less of a Cas9 (*e.g.*, the Cas9 of SEQ ID NO: 2). In some embodiments, the C-terminal portion that is rearranged to the N-terminus includes or corresponds to the C-terminal 410, 400, 390, 380, 370, 360, 350, 340, 330, 320, 310, 300, 290, 280, 270, 260, 250, 240, 230, 220, 210, 200, 190, 180, 170, 160, 150, 140, 130, 120, 110, 100, 90, 80, 70, 60, 50, 40, 30, 20, or 10 residues of a Cas9 (*e.g.*, the Cas9 of SEQ ID NO: 2). In some embodiments, the C-terminal portion that is rearranged to the N-terminus includes or corresponds to the C-terminal 357, 341, 328, 120, or 69 residues of a Cas9 (e.g., the Cas9 of SEQ ID NO: 2).

[370] In other embodiments, circular permutant Cas9 variants may be defined as a topological rearrangement of a Cas9 primary structure based on the following method, which is based on *S. pyogenes* Cas9 of SEQ ID NO: 2: (a) selecting a circular permutant (CP) site corresponding to an internal amino acid residue of the Cas9 primary structure, which dissects the original protein into two halves: an N-terminal region and a C-terminal region; (b) modifying the Cas9 protein sequence (e.g., by genetic engineering techniques) by moving the original C-terminal region (comprising the CP site amino acid) to precede the original N-terminal region, thereby forming a new N-terminus of the Cas9 protein that now begins with the CP site amino acid residue. The CP site can be located in any domain of the Cas9 protein, including, for example, the helical-II domain, the RuvCIII domain, or the CTD domain. For example, the CP site may be located (relative the *S. pyogenes* Cas9 of SEQ ID NO: 2) at original amino acid residue 181, 199, 230, 270, 310, 1010, 1016, 1023, 1029, 1041, 1247, 1249, or 1282. Thus, once relocated to the N-

terminus, original amino acid 181, 199, 230, 270, 310, 1010, 1016, 1023, 1029, 1041, 1247, 1249, or 1282 would become the new N-terminal amino acid. Nomenclature of these CP-Cas9 proteins may be referred to as Cas9-CP^{181}, Cas9-CP^{199}, Cas9-CP^{230}, Cas9-CP^{270}, Cas9-CP^{310}, Cas9-CP^{1010}, Cas9-CP^{1016}, Cas9-CP^{1023}, Cas9-CP^{1029}, Cas9-CP^{1041}, Cas9-CP^{1247}, Cas9-CP^{1249}, and Cas9-CP^{1282}, respectively. This description is not meant to be limited to making CP variants from SEQ ID NO: 2, but may be implemented to make CP variants in any Cas9 sequence, either at CP sites that correspond to these positions, or at other CP sites entirely. This description is not meant to limit the specific CP sites in any way. Virtually any CP site may be used to form a CP-Cas9 variant.

[371] Exemplary CP-Cas9 amino acid sequences, based on the Cas9 of SEQ ID NO: 2, are provided below in which linker sequences are indicated by underlining and optional methionine (M) residues are indicated in bold. It should be appreciated that the disclosure provides CP-Cas9 sequences that do not include a linker sequence or that include different linker sequences. It should be appreciated that CP-Cas9 sequences may be based on Cas9 sequences other than that of SEQ ID NO: 2 and any examples provided herein are not meant to be limiting. Exemplary CP-Cas9 sequences are as follows:

CP name	Sequence	SEQ ID NO:
CP1012	DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGE IVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYK EVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQ AENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLG GDGGSGGSGGSGGSGGSGGSGGDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLG NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIY LALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSA RLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDD DLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDL TLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPK HSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKED YFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFED REMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSD GFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVEN TQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVRE INNYHHAHDAYLNAVVGTALIKKYPKLESEFVYG	54

CP name	Sequence	SEQ ID NO:
CP1028	EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGSGGSGGSGG SGGSGGSGGMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVE EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRG HFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLI AQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGD QYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQL PEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQ RTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKT YAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQ LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGR HKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLY LYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDN VPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQ ITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHA HDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQ	55
CP1041	NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKK LKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISE FSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDGGSGGSGGSGGSGGSGGSGGDKKY SIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRL KRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIV DEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDV DKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGN LIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKN GYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHL GELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITP WNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTE GMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDI QKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARE NQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY VDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMK NYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSR MNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGT ALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYS	56
CP1249	PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGG DGGSGGSGGSGGSGGSGGSGGMDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLG NTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL	57

CP name	Sequence	SEQ ID NO:
	IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTY DDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEEL LVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPY YVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKV LPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQL KEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF LKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTV KVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEH PVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKV LTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQ FYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQE IGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLV VAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGS	
CP1300	KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDGGSGGSGGSGGSGGSGGSGGDKKYSIGLAIGTNSVGWAVITDEYKVPSK KFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNE MAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTD KADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQ LSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKR YDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIE KILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAI KKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKE LGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLK DDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAE RGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLV SDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVR KMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGR DFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS PTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNE QKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRD	58

[372] The Cas9 circular permutants may be useful in the prime editing constructs utilized in the methods and compositions described herein. Exemplary C-terminal fragments of Cas9, based on the Cas9 of SEQ ID NO: 2, which may be rearranged to an N-terminus of Cas9, are provided below. It should be appreciated that such C-terminal fragments of Cas9 are exemplary and are not meant to be limiting. These exemplary CP-Cas9 fragments have the following sequences:

CP name	Sequence	SEQ ID NO:
CP1012 C-terminal fragment	DYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETG EIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKG YKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEK LKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIR EQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ LGGD	59
CP1028 C-terminal fragment	EIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSL FELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	60
CP1041 C-terminal fragment	NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKK LKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISE FSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD	61
CP1249 C-terminal fragment	PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAE NIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGG D	62
CP1300 C-terminal fragment	KPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGD	63

I. Cas9 variants with modified PAM specificities

[373] The prime editors utilized in the methods and compositions of the present disclosure may also comprise Cas9 variants with modified PAM specificities. Some aspects of this disclosure provide Cas9 proteins that exhibit activity on a target sequence that does not comprise the canonical PAM (5′-NGG-3′, where N is A, C, G, or T) at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NGG-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NNG-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NNA-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NNC-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NNT-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NGT-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a

5′-NGA-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NGC-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NAA-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NAC-3′ PAM sequence at its 3′-end. In some embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NAT-3′ PAM sequence at its 3′-end. In still other embodiments, the Cas9 protein exhibits activity on a target sequence comprising a 5′-NAG-3′ PAM sequence at its 3′-end.

[374] It should be appreciated that any of the amino acid mutations described herein, (e.g., A262T) from a first amino acid residue (e.g., A) to a second amino acid residue (e.g., T) may also include mutations from the first amino acid residue to an amino acid residue that is similar to (e.g., conserved) the second amino acid residue. For example, mutation of an amino acid with a hydrophobic side chain (e.g., alanine, valine, isoleucine, leucine, methionine, phenylalanine, tyrosine, or tryptophan) may be a mutation to a second amino acid with a different hydrophobic side chain (e.g., alanine, valine, isoleucine, leucine, methionine, phenylalanine, tyrosine, or tryptophan). For example, a mutation of an alanine to a threonine (e.g., a A262T mutation) may also be a mutation from an alanine to an amino acid that is similar in size and chemical properties to a threonine, for example, serine. As another example, mutation of an amino acid with a positively charged side chain (e.g., arginine, histidine, or lysine) may be a mutation to a second amino acid with a different positively charged side chain (e.g., arginine, histidine, or lysine). As another example, mutation of an amino acid with a polar side chain (e.g., serine, threonine, asparagine, or glutamine) may be a mutation to a second amino acid with a different polar side chain (e.g., serine, threonine, asparagine, or glutamine). Additional similar amino acid pairs include, but are not limited to, the following: phenylalanine and tyrosine; asparagine and glutamine; methionine and cysteine; aspartic acid and glutamic acid; and arginine and lysine. The skilled artisan would recognize that such conservative amino acid substitutions will likely have minor effects on protein structure and are likely to be well tolerated without compromising function. In some embodiments, any amino of the amino acid mutations provided herein from one amino acid to a threonine may be an amino acid mutation to a serine. In some embodiments, any amino of the amino acid mutations provided herein from one amino acid to an arginine may be an amino acid mutation to a lysine. In some embodiments, any amino of the amino acid

mutations provided herein from one amino acid to an isoleucine, may be an amino acid mutation to an alanine, valine, methionine, or leucine. In some embodiments, any amino of the amino acid mutations provided herein from one amino acid to a lysine may be an amino acid mutation to an arginine. In some embodiments, any amino of the amino acid mutations provided herein from one amino acid to an aspartic acid may be an amino acid mutation to a glutamic acid or asparagine. In some embodiments, any amino of the amino acid mutations provided herein from one amino acid to a valine may be an amino acid mutation to an alanine, isoleucine, methionine, or leucine. In some embodiments, any amino of the amino acid mutations provided herein from one amino acid to a glycine may be an amino acid mutation to an alanine. It should be appreciated, however, that additional conserved amino acid residues would be recognized by the skilled artisan and any of the amino acid mutations to other conserved amino acid residues are also within the scope of this disclosure.

[375] In some embodiments, the Cas9 protein comprises a combination of mutations that exhibit activity on a target sequence comprising a 5′-NAA-3′ PAM sequence at its 3′-end. In some embodiments, the combination of mutations are present in any one of the clones listed in Table 1. In some embodiments, the combination of mutations are conservative mutations of the clones listed in Table 1. In some embodiments, the Cas9 protein comprises the combination of mutations of any one of the Cas9 clones listed in Table 1.

[376] Table 1: NAA PAM Clones

Mutations from wild-type SpCas9 (e.g., SEQ ID NO: 2)
D177N, K218R, D614N, D1135N, P1137S, E1219V, A1320V, A1323D, R1333K
D177N, K218R, D614N, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, G715C, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A367T, K710E, R1114G, D1135N, P1137S, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, D861N, D1135N, K1188R, E1219V, Q1221H, H1264H, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, V743I, R753G, M1021T, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, V743I, R753G, E762G, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, S1274R, A1320V, R1333K
A10T, I322V, S409I, E427G, A589S, R753G, D1135N, E1219V, Q1221H, H1264H, A1320V, R1333K
A10T, I322V, S409I, E427G, R753G, E757K, G865G, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, R753G, E757K, D1135N, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, K599R, M631A, R654L, K673E, V743I, R753G, N758H, E762G, D1135N, D1180G, E1219V, Q1221H, Q1256R, H1264Y, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N869S, N1054D, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, A1320V, A1323D, R1333K

A10T, I322V, S409I, E427G, R654L, L727I, V743I, R753G, E762G, R859S, N946D, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, N1317T, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, K1151E, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, V1290G, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, N803S, N869S, L921P, Y1016D, G1077D, F1080S, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, E630K, R654L, K673E, V743I, R753G, E762G, Q768H, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, Q768H, N803S, N869S, Y1016D, G1077D, R1114G, F1134L, D1135N, D1180G, E1219V, Q1221H, G1223S, H1264Y, L1318S, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, F693L, V743I, R753G, E762G, N803S, N869S, L921P, Y1016D, G1077D, F1801S, R1114G, D1135N, D1180G, E1219V, Q1221H, H1264Y, L1318S, A1320V, A1323D, R1333K
A10T, I322V, S409I, E427G, R654L, V743I, R753G, M1021T, D1135N, D1180G, K1211R, E1219V, Q1221H, H1264Y, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, M673I, N803S, N869S, G1077D, R1114G, D1135N, V1139A, D1180G, E1219V, Q1221H, A1320V, R1333K
A10T, I322V, S409I, E427G, R654L, K673E, V743I, R753G, E762G, N803S, N869S, R1114G, D1135N, E1219V, Q1221H, A1320V, R1333K

[377] In some embodiments, the Cas9 protein comprises an amino acid sequence that is at least 80% identical to the amino acid sequence of a Cas9 protein as provided by any one of the variants of Table 1. In some embodiments, the Cas9 protein comprises an amino acid sequence that is at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence of a Cas9 protein as provided by any one of the variants of Table 1.

[378] In some embodiments, the Cas9 protein exhibits an increased activity on a target sequence that does not comprise the canonical PAM (5′-NGG-3′) at its 3′ end as compared to *Streptococcus pyogenes* Cas9 as provided by SEQ ID NO: 2. In some embodiments, the Cas9 protein exhibits an activity on a target sequence having a 3′ end that is not directly adjacent to the canonical PAM sequence (5′-NGG-3′) that is at least 5-fold increased as compared to the activity of *Streptococcus pyogenes* Cas9 as provided by SEQ ID NO: 2 on the same target sequence. In some embodiments, the Cas9 protein exhibits an activity on a target sequence that is not directly adjacent to the canonical PAM sequence (5′-NGG-3′) that is at least 10-fold, at least 50-fold, at least 100-fold, at least 500-fold, at least 1,000-fold, at least 5,000-fold, at least

10,000-fold, at least 50,000-fold, at least 100,000-fold, at least 500,000-fold, or at least 1,000,000-fold increased as compared to the activity of *Streptococcus pyogenes* as provided by SEQ ID NO: 2 on the same target sequence. In some embodiments, the 3′ end of the target sequence is directly adjacent to an AAA, GAA, CAA, or TAA sequence. In some embodiments, the Cas9 protein comprises a combination of mutations that exhibit activity on a target sequence comprising a 5′-NAC-3′ PAM sequence at its 3′-end. In some embodiments, the combination of mutations are present in any one of the clones listed in Table 2. In some embodiments, the combination of mutations are conservative mutations of the clones listed in Table 2. In some embodiments, the Cas9 protein comprises the combination of mutations of any one of the Cas9 clones listed in Table 2.

[379] Table 2: NAC PAM Clones

Mutations from wild-type SpCas9 (e.g., SEQ ID NO: 2)
T472I, R753G, K890E, D1332N, R1335Q, T1337N
I1057S, D1135N, P1301S, R1335Q, T1337N
T472I, R753G, D1332N, R1335Q, T1337N
D1135N, E1219V, D1332N, R1335Q, T1337N
T472I, R753G, K890E, D1332N, R1335Q, T1337N
I1057S, D1135N, P1301S, R1335Q, T1337N
T472I, R753G, D1332N, R1335Q, T1337N
T472I, R753G, Q771H, D1332N, R1335Q, T1337N
E627K, T638P, K652T, R753G, N803S, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, T638P, K652T, R753G, N803S, K959N, R1114G, D1135N, K1156E, E1219V, D1332N, R1335Q, T1337N
E627K, T638P, V647I, R753G, N803S, K959N, G1030R, I1055E, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, E630G, T638P, V647A, G687R, N767D, N803S, K959N, R1114G, D1135N, E1219V, D1332G, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, R1114G, D1135N, E1219V, N1266H, D1332N, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, I1057T, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, T638P, R753G, N803S, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
E627K, M631I, T638P, R753G, N803S, K959N, Y1036H, R1114G, D1135N, E1219V, D1251G, D1332G, R1335Q, T1337N
E627K, T638P, R753G, N803S, V875I, K959N, Y1016C, R1114G, D1135N, E1219V, D1251G, D1332G, R1335Q, T1337N, I1348V
K608R, E627K, T638P, V647I, R654L, R753G, N803S, T804A, K848N, V922A, K959N, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, R753G, N803S, V922A, K959N, K1014N, V1015A, R1114G, D1135N, K1156N, E1219V, N1252D, D1332N, R1335Q, T1337N
K608R, E627K, R629G, T638P, V647I, A711T, R753G, K775R, K789E, N803S, K959N, V1015A, Y1036H, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, R753G, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N

I670S, K608R, E627K, E630G, T638P, V647I, R653K, R753G, I795L, K797N, N803S, K866R, K890N, K959N, Y1016C, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, T740A, G752R, R753G, K797N, N803S, K948E, K959N, V1015A, Y1016S, R1114G, D1135N, E1219V, N1266H, D1332N, R1335Q, T1337N
I570T, A589V, K608R, E627K, T638P, V647I, R654L, Q716R, R753G, N803S, K948E, K959N, Y1016S, R1114G, D1135N, E1207G, E1219V, N1234D, D1332N, R1335Q, T1337N
K608R, E627K, R629G, T638P, V647I, R654L, Q740R, R753G, N803S, K959N, N990S, T995S, V1015A, Y1036D, R1114G, D1135N, E1207G, E1219V, N1234D, N1266H, D1332N, R1335Q, T1337N
I562F, V565D, I570T, K608R, L625S, E627K, T638P, V647I, R654I, G752R, R753G, N803S, N808D, K959N, M1021L, R1114G, D1135N, N1177S, N1234D, D1332N, R1335Q, T1337N
I562F, I570T, K608R, E627K, T638P, V647I, R753G, E790A, N803S, K959N, V1015A, Y1036H, R1114G, D1135N, D1180E, A1184T, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, E627K, T638P, V647I, R654H, R753G, E790A, N803S, K959N, V1015A, R1114G, D1127A, D1135N, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, L625S, E627K, T638P, V647I, R654I, T703P, R753G, N803S, N808D, K959N, M1021L, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
I570S, K608R, E627K, E630G, T638P, V647I, R653K, R753G, I795L, N803S, K866R, K890N, K959N, Y1016C, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N
I570T, K608R, E627K, T638P, V647I, R654H, R753G, E790A, N803S, K959N, V1016A, R1114G, D1135N, E1219V, K1246E, D1332N, R1335Q, T1337N
K608R, E627K, T638P, V647I, R654L, K673E, R753G, E790A, N803S, K948E, K959N, R1114G, D1127G, D1135N, D1180E, E1219V, N1286H, D1332N, R1335Q, T1337N
K608R, L625S, E627K, T638P, V647I, R654I, I670T, R753G, N803S, N808D, K959N, M1021L, R1114G, D1135N, E1219V, N1286H, D1332N, R1335Q, T1337N
E627K, M631V, T638P, V647I, K710E, R753G, N803S, N808D, K948E, M1021L, R1114G, D1135N, E1219V, D1332N, R1335Q, T1337N, S1338T, H1349R

[380] In some embodiments, the Cas9 protein comprises an amino acid sequence that is at least 80% identical to the amino acid sequence of a Cas9 protein as provided by any one of the variants of Table 2. In some embodiments, the Cas9 protein comprises an amino acid sequence that is at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence of a Cas9 protein as provided by any one of the variants of Table 2.

[381] In some embodiments, the Cas9 protein exhibits an increased activity on a target sequence that does not comprise the canonical PAM (5′-NGG-3′) at its 3′ end as compared to *Streptococcus pyogenes* Cas9 as provided by SEQ ID NO: 2. In some embodiments, the Cas9 protein exhibits an activity on a target sequence having a 3′ end that is not directly adjacent to the canonical PAM sequence (5′-NGG-3′) that is at least 5-fold increased as compared to the activity of *Streptococcus pyogenes* Cas9 as provided by SEQ ID NO: 2 on the same target sequence. In some embodiments, the Cas9 protein exhibits an activity on a target sequence that is not directly adjacent to the canonical PAM sequence (5′-NGG-3′) that is at least 10-fold, at least 50-fold, at least 100-fold, at least 500-fold, at least 1,000-fold, at least 5,000-fold, at least 10,000-fold, at least 50,000-fold, at least 100,000-fold, at least 500,000-fold, or at least

1,000,000-fold increased as compared to the activity of *Streptococcus pyogenes* as provided by SEQ ID NO: 2 on the same target sequence. In some embodiments, the 3′ end of the target sequence is directly adjacent to an AAC, GAC, CAC, or TAC sequence.

[382] In some embodiments, the Cas9 protein comprises a combination of mutations that exhibit activity on a target sequence comprising a 5′-NAT-3′ PAM sequence at its 3′-end. In some embodiments, the combination of mutations are present in any one of the clones listed in Table 3. In some embodiments, the combination of mutations are conservative mutations of the clones listed in Table 3. In some embodiments, the Cas9 protein comprises the combination of mutations of any one of the Cas9 clones listed in Table 3.

[383] Table 3: NAT PAM Clones

Mutations from wild-type SpCas9 (e.g., SEQ ID NO: 2)
K961E, H985Y, D1135N, K1191N, E1219V, Q1221H, A1320A, P1321S, R1335L
D1135N, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
V743I, R753G, E790A, D1135N, G1218S, E1219V, Q1221H, A1227V, P1249S, N1286K, A1293T, P1321S, D1322G, R1335L, T1339I
F575S, M631L, R654L, V748I, V743I, R753G, D853E, V922A, R1114G D1135N, G1218S, E1219V, Q1221H, A1227V, P1249S, N1286K, A1293T, P1321S, D1322G, R1335L, T1339I
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, N1286K, P1321S, D1322G, R1335L
M631L, R654L, R753G, K797E, D853E, V922A, D1012A, R1114G D1135N, G1218S, E1219V, Q1221H, P1249S, N1317K, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, D596Y, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, Q1256R, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, K710E, V750A, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, K649R, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, K1156E, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1322G, R1335L
F575S, M631L, R654L, R664K, R753G, D853E, V922A, I1057G, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, N1308D, P1321S, D1322G, R1335L
M631L, R654L, R753G, D853E, V922A, R1114G, Y1131C, D1135N, E1150V, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L
M631L, R654L, R664K, R753G, D853E, I1057V, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L
M631L, R654L, R664K, R753G, I1057V, R1114G, Y1131C, D1135N, D1180G, G1218S, E1219V, Q1221H, P1249S, P1321S, D1332G, R1335L

[384] The above description of various napDNAbps which can be used in connection with the prime editors is not meant to be limiting in any way. The prime editors may comprise the canonical SpCas9, or any ortholog Cas9 protein, or any variant Cas9 protein—including any

naturally occurring variant, mutant, or otherwise engineered version of Cas9—that is known or which can be made or evolved through a directed evolutionary or otherwise mutagenic process. In various embodiments, the Cas9 or Cas9 variants have a nickase activity, i.e., only cleave of strand of the target DNA sequence. In other embodiments, the Cas9 or Cas9 variants have inactive nucleases, i.e., are "dead" Cas9 proteins. Other variant Cas9 proteins that may be used are those having a smaller molecular weight than the canonical SpCas9 (e.g., for easier delivery) or having modified or rearranged primary amino acid structure (e.g., the circular permutant formats). The prime editors utilized in the methods and compositions described herein may also comprise Cas9 equivalents, including Cas12a/Cpf1 and Cas12b proteins which are the result of convergent evolution. The napDNAbps used herein (e.g., SpCas9, Cas9 variant, or Cas9 equivalents) may also contain various modifications that alter/enhance their PAM specificities. Lastly, the application contemplates any Cas9, Cas9 variant, or Cas9 equivalent which has at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.9% sequence identity to a reference Cas9 sequence, such as a references SpCas9 canonical sequences or a reference Cas9 equivalent (e.g., Cas12a/Cpf1).

In a particular embodiment, the Cas9 variant having expanded PAM capabilities is SpCas9 (H840A) VRQR (SEQ ID NO: 64), which has the following amino acid sequence (with the V, R, Q, R substitutions relative to the SpCas9 (H840A) of SEQ ID NO: 33 being show in bold underline. In addition, the methionine residue in SpCas9 (H840) was removed for SpCas9 (H840A) VRQR):

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKL
VDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS
ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD
QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR
RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER
MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK
QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLT
FKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREIN
NYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK
TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK
LIARKKDWDPKKYGGFVSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY

KEVKKDLIIKLPKYSLFELENGRKRMLASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD
TTIDRK**Q**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 64)

[385] In another particular embodiment, the Cas9 variant having expanded PAM capabilities is SpCas9 (H840A) VRER, which has the following amino acid sequence (with the V, R, E, R substitutions relative to the SpCas9 (H840A) of SEQ ID NO: 51 being shown in bold underline . In addition, the methionine residue in SpCas9 (H840) was removed for SpCas9 (H840A) VRER):

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRR
YTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKL
VDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILS
ARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG
DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFD
QSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILR
RQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIER
MTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK
QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE
RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLT
FKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQK
GQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAI
VPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLS
ELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREIN
NYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFK
TEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDK
LIARKKDWDPKKYGGF**V**SPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGY
KEVKKDLIIKLPKYSLFELENGRKRMLASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ
LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFD
TTIDRK**E**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 65)

[386] In some embodiments, the napDNAbp that functions with a non-canonical PAM sequence is an Argonaute protein. One example of such a nucleic acid programmable DNA binding protein is an Argonaute protein from *Natronobacterium gregoryi* (NgAgo). NgAgo is a ssDNA-guided endonuclease. NgAgo binds 5′ phosphorylated ssDNA of ~24 nucleotides (gDNA) to guide it to its target site and will make DNA double-strand breaks at the gDNA site. In contrast to Cas9, the NgAgo–gDNA system does not require a protospacer-adjacent motif (PAM). Using a nuclease inactive NgAgo (dNgAgo) can greatly expand the bases that may be targeted. The characterization and use of NgAgo have been described in Gao *et al.*, *Nat Biotechnol.*, 2016 Jul;34(7):768-73. PubMed PMID: 27136078; Swarts *et al.*, *Nature*. 507(7491) (2014):258-61; and Swarts *et al.*, *Nucleic Acids Res.* 43(10) (2015):5120-9, each of which is incorporated herein by reference.

[387] In some embodiments, the napDNAbp is a prokaryotic homolog of an Argonaute protein. Prokaryotic homologs of Argonaute proteins are known and have been described, for example, in Makarova K., *et al.*, "Prokaryotic homologs of Argonaute proteins are predicted to function as key components of a novel system of defense against mobile genetic elements", *Biol Direct.* 2009 Aug 25;4:29. doi: 10.1186/1745-6150-4-29, the entire contents of which is hereby incorporated by reference. In some embodiments, the napDNAbp is a Marinitoga piezophila Argonaute (MpAgo) protein. The CRISPR-associated Marinitoga piezophila Argonaute (MpAgo) protein cleaves single-stranded target sequences using 5′-phosphorylated guides. The 5′ guides are used by all known Argonautes. The crystal structure of an MpAgo-RNA complex shows a guide strand binding site comprising residues that block 5′ phosphate interactions. This data suggests the evolution of an Argonaute subclass with noncanonical specificity for a 5′-hydroxylated guide. See, e.g., Kaya *et al.*, "A bacterial Argonaute with noncanonical guide RNA specificity", *Proc Natl Acad Sci U S A.* 2016 Apr 12;113(15):4057-62, the entire contents of which are hereby incorporated by reference). It should be appreciated that other argonaute proteins may be used, and are within the scope of this disclosure.

[388] Some aspects of the disclosure provide Cas9 domains that have different PAM specificities. Typically, Cas9 proteins, such as Cas9 from *S. pyogenes* (spCas9), require a canonical NGG PAM sequence to bind a particular nucleic acid region. This may limit the ability to edit desired bases within a genome. In some embodiments, the base editing fusion proteins provided herein may need to be placed at a precise location, for example where a target base is placed within a 4 base region (*e.g.*, a "editing window"), which is approximately 15 bases upstream of the PAM. See Komor, A.C., *et al.*, "Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage" *Nature* 533, 420-424 (2016), the entire contents of which are hereby incorporated by reference. Accordingly, in some embodiments, any of the fusion proteins provided herein may contain a Cas9 domain that is capable of binding a nucleotide sequence that does not contain a canonical (*e.g.*, NGG) PAM sequence. Cas9 domains that bind to non-canonical PAM sequences have been described in the art and would be apparent to the skilled artisan. For example, Cas9 domains that bind non-canonical PAM sequences have been described in Kleinstiver, B. P., *et al.*, "Engineered CRISPR-Cas9 nucleases with altered PAM specificities" *Nature* 523, 481-485 (2015); and Kleinstiver, B. P., *et al.*, "Broadening the targeting range of *Staphylococcus aureus* CRISPR-Cas9 by modifying PAM recognition" *Nature*

Biotechnology 33, 1293-1298 (2015); the entire contents of each are hereby incorporated by reference.

[389] For example, a napDNAbp domain with altered PAM specificity, such as a domain with at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity with wild type *Francisella novicida* Cpf1 (D917, E1006, and D1255) (SEQ ID NO: 66), which has the following amino acid sequence:

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYHQFFIEEILSSVCISE
DLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSEKFKNLFNQNLIDAKKGQESDLILWL
KQSKDNGIELFKANSDITDIDEALEIIKSFKGWTTYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLEN
KAKYESLKDKAPEAINYEQIKKDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIG
GKFVNGENTKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTMQSFY
EQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDYSVIGTAVLEYITQQIAP
KNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDIDKQCRFEEILANFAAIPMIFDEIAQNKDN
LAQISIKYQNQGKKDLLQASAEDDVKAIKDLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEEC
YFELANIVPLYNKIRNYITQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKN
NKIFDDKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKNGSPQKGY
EKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVENQGYKLTFENISESYIDSVVN
QGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDERNLQDVVYKLNGEAELFYRKQSIPKKITHPAKEA
IANKNKDNPKKESVFEYDLIKDKRFTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSIDRGE
RHLAYYTLVDGKGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVFEDLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEFDKTGGVLRAYQL
TAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESVSKSQEFFSKFDKICYNLDKGYFEFSF
DYKNFGDKAAKGKWTIASFGSRLINFRNSDKNHNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICG
ESDKKFFAKLTSVLNTILQMRNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDADANGAYHIGLKG
LMLLGRIKNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 66)

[390] An additional napDNAbp domain with altered PAM specificity, such as a domain having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity with wild type *Geobacillus thermodenitrificans* Cas9 (SEQ ID NO: 20), which has the following amino acid sequence:

MKYKIGLDIGITSIGWAVINLDIPRIEDLGVRIFDRAENPKTGESLALPRRLARSARRRLRRRKHRLERIRR
LFVREGILTKEELNKLFEKKHEIDVWQLRVEALDRKLNNDELARILLHLAKRRGFRSNRKSERTNKENST
MLKHIEENQSILSSYRTVAEMVVKDPKFSLHKRNKEDNYTNTVARDDLEREIKLIFAKQREYGNIVCTEA
FEHEYISIWASQRPFASKDDIEKKVGFCTFEPKEKRAPKATYTFQSFTVWEHINKLRLVSPGGIRALTDDE
RRLIYKQAFHKNKITFHDVRTLLNLPDDTRFKGLLYDRNTTLKENEKVRFLELGAYHKIRKAIDSVYGK
GAAKSFRPIDFDTFGYALTMFKDDTDIRSYLRNEYEQNGKRMENLADKVYDEELIEELLNLSFSKFGHLS
LKALRNILPYMEQGEVYSTACERAGYTFTGPKKKQKTVLLPNIPPIANPVVMRALTQARKVVNAIIKKY
GSPVSIHIELARELSQSFDERRKMQKEQEGNRKKNETAIRQLVEYGLTLNPTGLDIVKFKLWSEQNGKCA
YSLQPIEIERLLEPGYTEVDHVIPYSRSLDDSYTNKVLVLTKENREKGNRTPAEYLGLGSERWQQFETFVL
TNKQFSKKKRDRLLRLHYDENEENEFKNRNLNDTRYISRFLANFIREHLKFADSDDKQKVYTVNGRITA
HLRSRWNFNKNREESNLHHAVDAAIVACTTPSDIARVTAFYQRREQNKELSKKTDPQFPQPWPHFADEL
QARLSKNPKESIKALNLGNYDNEKLESLQPVFVSRMPKRSITGAAHQETLRRYIGIDERSGKIQTVVKKK
LSEIQLDKTGHFPMYGKESDPRTYEAIRQRLLEHNNDPKKAFQEPLYKPKKNGELGPIIRTIKIIDTTNQVI
PLNDGKTVAYNSNIVRVDVFEKDGKYYCVPIYTIDMMKGILPNKAIEPNKPYSEWKEMTEDYTFRFSLY
PNDLIRIEFPREKTIKTAVGEEIKIKDLFAYYQTIDSSNGGLSLVSHDNNFSLRSIGSRTLKRFEKYQVDVL
GNIYKVRGEKRVGVASSSHSKAGETIRPL (SEQ ID NO: 20)

[391] In some embodiments, the nucleic acid programmable DNA binding protein (napDNAbp) is a nucleic acid programmable DNA binding protein that does not require a canonical (NGG) PAM sequence. In some embodiments, the napDNAbp is an argonaute protein. One example of such a nucleic acid programmable DNA binding protein is an Argonaute protein from *Natronobacterium gregoryi* (NgAgo). NgAgo is a ssDNA-guided endonuclease. NgAgo binds 5′ phosphorylated ssDNA of ~24 nucleotides (gDNA) to guide it to its target site and will make DNA double-strand breaks at the gDNA site. In contrast to Cas9, the NgAgo–gDNA system does not require a protospacer-adjacent motif (PAM). Using a nuclease inactive NgAgo (dNgAgo) can greatly expand the bases that may be targeted. The characterization and use of NgAgo have been described in Gao *et al.*, *Nat Biotechnol.*, 34(7): 768-73 (2016), PubMed PMID: 27136078; Swarts *et al.*, *Nature*, 507(7491): 258-61 (2014); and Swarts *et al.*, *Nucleic Acids Res.* 43(10) (2015): 5120-9, each of which is incorporated herein by reference. The sequence of *Natronobacterium gregoryi* Argonaute is provided in SEQ ID NO: 67.

[392] The disclosed fusion proteins may comprise a napDNAbp domain having at least 80%, at least 85%, at least 90%, at least 95%, or at least 99% sequence identity with wild type *Natronobacterium gregoryi* Argonaute (SEQ ID NO: 67), which has the following amino acid sequence:

MTVIDLDSTTTADELTSGHTYDISVTLTGVYDNTDEQHPRMSLAFEQDNGERRYITLWKNTTPKDVFTY
DYATGSTYIFTNIDYEVKDGYENLTATYQTTVENATAQEVGTTDEDETFAGGEPLDHHLDDALNETPDD
AETESDSGHVMTSFASRDQLPEWTLHTYTLTATDGAKTDTEYARRTLAYTVRQELYTDHDAAPVATDG
LMLLTPEPLGETPLDLDCGVRVEADETRTLDYTTAKDRLLARELVEEGLKRSLWDDYLVRGIDEVLSKE
PVLTCDEFDLHERYDLSVEVGHSGRAYLHINFRHRFVPKLTLADIDDDNIYPGLRVKTTYRPRRGHIVWG
LRDECATDSLNTLGNQSVVAYHRNNQTPINTDLLDAIEAADRRVVETRRQGHGDDAVSFPQELLAVEPN
THQIKQFASDGFHQQARSKTRLSASRCSEKAQAFAERLDPVRLNGSTVEFSSEFFTGNNEQQLRLLYENG
ESVLTFRDGARGAHPDETFSKGIVNPPESFEVAVVLPEQQADTCKAQWDTMADLLNQAGAPPTRSETVQ
YDAFSSPESISLNVAGAIDPSEVDAAFVVLPPDQEGFADLASPTETYDELKKALANMGIYSQMAYFDRFR
DAKIFYTRNVALGLLAAAGGVAFTTEHAMPGDADMFIGIDVSRSYPEDGASGQINIAATATAVYKDGTI
LGHSSTRPQLGEKLQSTDVRDIMKNAILGYQQVTGESPTHIVIHRDGFMNEDLDPATEFLNEQGVEYDIV
EIRKQPQTRLLAVSDVQYDTPVKSIAAINQNEPRATVATFGAPEYLATRDGGGLPRPIQIERVAGETDIET
LTRQVYLLSQSHIQVHNSTARLPITTAYADQASTHATKGYLVQTGAFESNVGFL (SEQ ID NO: 67)

[393] In addition, any available methods may be utilized to obtain or construct a variant or mutant Cas9 protein. The term "mutation," as used herein, refers to a substitution of a residue within a sequence, e.g., a nucleic acid or amino acid sequence, with another residue, or a deletion or insertion of one or more residues within a sequence. Mutations are typically described herein by identifying the original residue followed by the position of the residue within the sequence and by the identity of the newly substituted residue. Various methods for making the amino acid

substitutions (mutations) provided herein are well known in the art, and are provided by, for example, Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)). Mutations can include a variety of categories, such as single base polymorphisms, microduplication regions, indel, and inversions, and is not meant to be limiting in any way. Mutations can include "loss-of-function" mutations which is the normal result of a mutation that reduces or abolishes a protein activity. Most loss-of-function mutations are recessive, because in a heterozygote the second chromosome copy carries an unmutated version of the gene coding for a fully functional protein whose presence compensates for the effect of the mutation. Mutations also embrace "gain-of-function" mutations, which is one which confers an abnormal activity on a protein or cell that is otherwise not present in a normal condition. Many gain-of-function mutations are in regulatory sequences rather than in coding regions, and can therefore have a number of consequences. For example, a mutation might lead to one or more genes being expressed in the wrong tissues, these tissues gaining functions that they normally lack. Because of their nature, gain-of-function mutations are usually dominant.

[394] Mutations can be introduced into a reference Cas9 protein using site-directed mutagenesis. Older methods of site-directed mutagenesis known in the art rely on sub-cloning of the sequence to be mutated into a vector, such as an M13 bacteriophage vector, that allows the isolation of single-stranded DNA template. In these methods, one anneals a mutagenic primer (i.e., a primer capable of annealing to the site to be mutated but bearing one or more mismatched nucleotides at the site to be mutated) to the single-stranded template and then polymerizes the complement of the template starting from the 3′ end of the mutagenic primer. The resulting duplexes are then transformed into host bacteria and plaques are screened for the desired mutation. More recently, site-directed mutagenesis has employed PCR methodologies, which have the advantage of not requiring a single-stranded template. In addition, methods have been developed that do not require sub-cloning. Several issues must be considered when PCR-based site-directed mutagenesis is performed. First, in these methods it is desirable to reduce the number of PCR cycles to prevent expansion of undesired mutations introduced by the polymerase. Second, a selection must be employed in order to reduce the number of non-mutated parental molecules persisting in the reaction. Third, an extended-length PCR method is preferred in order to allow the use of a single PCR primer set. And fourth, because of the non-template-

dependent terminal extension activity of some thermostable polymerases it is often necessary to incorporate an end-polishing step into the procedure prior to blunt-end ligation of the PCR-generated mutant product.

[395] Mutations may also be introduced by directed evolution processes, such as phage-assisted continuous evolution (PACE) or phage-assisted noncontinuous evolution (PANCE). The term "phage-assisted continuous evolution (PACE)," as used herein, refers to continuous evolution that employs phage as viral vectors. The general concept of PACE technology has been described, for example, in International PCT Application, PCT/US2009/056194, filed September 8, 2009, published as WO 2010/028347 on March 11, 2010; International PCT Application, PCT/US2011/066747, filed December 22, 2011, published as WO 2012/088381 on June 28, 2012; U.S. Application, U.S. Patent No. 9,023,594, issued May 5, 2015, International PCT Application, PCT/US2015/012022, filed January 20, 2015, published as WO 2015/134121 on September 11, 2015, and International PCT Application, PCT/US2016/027795, filed April 15, 2016, published as WO 2016/168631 on October 20, 2016, the entire contents of each of which are incorporated herein by reference. Variant Cas9s may also be obtain by phage-assisted non-continuous evolution (PANCE)," which as used herein, refers to non-continuous evolution that employs phage as viral vectors. PANCE is a simplified technique for rapid in vivo directed evolution using serial flask transfers of evolving 'selection phage' (SP), which contain a gene of interest to be evolved, across fresh *E. coli* host cells, thereby allowing genes inside the host *E. coli* to be held constant while genes contained in the SP continuously evolve. Serial flask transfers have long served as a widely-accessible approach for laboratory evolution of microbes, and, more recently, analogous approaches have been developed for bacteriophage evolution. The PANCE system features lower stringency than the PACE system.

[396] Any of the references noted above which relate to Cas9 or Cas9 equivalents are hereby incorporated by reference in their entireties, if not already stated so.

Other programmable nucleases

[397] In various embodiments described herein, the prime editors comprise a napDNAbp, such as a Cas9 protein. These proteins are "programmable" by way of their becoming complexed with a guide RNA (or a PEgRNA, as the case may be), which guides the Cas9 protein to a target site on the DNA which possess a sequence that is complementary to the spacer portion of the gRNA (or PEgRNA) and also which possesses the required PAM sequence. However, in certain

embodiment envisioned here, the napDNAbp may be substituted with a different type of programmable protein, such as a zinc finger nuclease or a transcription activator-like effector nuclease (TALEN).

[398] It is contemplated that suitable nucleases do not necessarily need to be "programmed" by a nucleic acid targeting molecule (such as a guide RNA), but rather, may be programmed by defining the specificity of a DNA-binding domain, such as and in particular, a nuclease. Just as in prime editing with napDNAbp moieties, it is preferable that such alternative programmable nucleases be modified such that only one strand of a target DNA is cut. In other words, the programmable nucleases should function as nickases, preferably. Once a programmable nuclease is selected (e.g., a ZFN or a TALEN), then additional functionalities may be engineered into the system to allow it to operate in accordance with a prime editing-like mechanism. For example, the programmable nucleases may be modified by coupling (e.g., via a chemical linker) an RNA or DNA extension arm thereto, wherein the extension arm comprises a primer binding site (PBS) and a DNA synthesis template. The programmable nuclease may also be coupled (e.g., via a chemical or amino acid linker) to a polymerase, the nature of which will depend upon whether the extension arm is DNA or RNA. In the case of an RNA extension arm, the polymerase can be an RNA-dependent DNA polymerase (e.g., reverse transcriptase). In the case of a DNA extension arm, the polymerase can be a DNA-dependent DNA polymerase (e.g., a prokaryotic polymerase, including Pol I, Pol II, or Pol III, or a eukaryotic polymerase, including Pol a, Pol b, Pol g, Pol d, Pol e, or Pol z). The system may also include other functionalities added as fusions to the programmable nucleases, or added *in trans* to facilitate the reaction as a whole (e.g., (a) a helicase to unwind the DNA at the cut site to make the cut strand with the 3′ end available as a primer, (b) a FEN1 to help remove the endogenous strand on the cut strand to drive the reaction towards replacement of the endogenous strand with the synthesized strand, or (c) a nCas9:gRNA complex to create a second site nick on the opposite strand, which may help drive the integration of the synthesize repair through favored cellular repair of the non-edited strand). In an analogous manner to prime editing with a napDNAbp, such a complex with an otherwise programmable nuclease could be used to synthesize and then install a newly synthesized replacement strand of DNA carrying an edit of interest permanently into a target site of DNA.

[399] Suitable alternative programmable nucleases are well known in the art which may be used in place of a napDNAbp:gRNA complex to construct an alternative prime editor system that

can be programmed to selectively bind a target site of DNA, and which can be further modified in the manner described above to co-localize a polymerase and an RNA or DNA extension arm comprising a primer binding site and a DNA synthesis template to specific nick site . For example, Transcription Activator-Like Effector Nucleases (TALENs) may be used as the programmable nuclease in the prime editing methods and compositions of matter described herein. TALENS are artificial restriction enzymes generated by fusing the TAL effector DNA binding domain to a DNA cleavage domain. These reagents enable efficient, programmable, and specific DNA cleavage and represent powerful tools for genome editing *in situ*. Transcription activator-like effectors (TALEs) can be quickly engineered to bind practically any DNA sequence. The term TALEN, as used herein, is broad and includes a monomeric TALEN that can cleave double stranded DNA without assistance from another TALEN. The term TALEN is also used to refer to one or both members of a pair of TALENs that are engineered to work together to cleave DNA at the same site. TALENs that work together may be referred to as a left-TALEN and a right-TALEN, which references the handedness of DNA. See U.S. Ser. No. 12/965,590; U.S. Ser. No. 13/426,991 (U.S. Pat. No. 8,450,471); U.S. Ser. No. 13/427,040 (U.S. Pat. No. 8,440,431); U.S. Ser. No. 13/427,137 (U.S. Pat. No. 8,440,432); and U.S. Ser. No. 13/738,381, all of which are incorporated by reference herein in their entirety. In addition, TALENS are described in WO 2015/027134, US 9,181,535, Boch *et al.*, "Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors", Science, vol. 326, pp. 1509-1512 (2009), Bogdanove *et al.*, TAL Effectors: Customizable Proteins for DNA Targeting, Science, vol. 333, pp. 1843-1846 (2011), Cade *et al.*, "Highly efficient generation of heritable zebrafish gene mutations using homo- and heterodimeric TALENs", Nucleic Acids Research, vol. 40, pp. 8001-8010 (2012), and Cermak *et al.*, "Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting", Nucleic Acids Research, vol. 39, No. 17, e82 (2011), each of which are incorporated herein by reference.

[400] Zinc finger nucleases may also be used as alternative programmable nucleases for use in the prime editing methods and compositions disclosed herein in place of napDNAbps, such as Cas9 nickases. Like with TALENS, the ZFN proteins may be modified such that they function as nickases, i.e., engineering the ZFN such that it cleaves only one strand of the target DNA in a manner similar to the napDNAbp used with the prime editors in the methods and compositions described herein. ZFN proteins have been extensively described in the art, for example, in

Carroll *et al.*, "Genome Engineering with Zinc-Finger Nucleases," *Genetics*, Aug 2011, Vol. 188: 773-782; Durai *et al.*, "Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells," *Nucleic Acids Res*, 2005, Vol. 33: 5978-90; and Gaj *et al.*, "ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering," *Trends Biotechnol.* 2013, Vol.31: 397-405, each of which are incorporated herein by reference in their entireties.

Prime editors: polymerase domain (e.g., reverse transcriptases)

[401] The present disclosure provides compositions and methods for prime editing with improved editing efficiency and/or reduced indel formation by inhibiting the DNA mismatch repair pathway while conducting prime editing of a target site. Accordingly, the present disclosure provides a method for editing a nucleic acid molecule by prime editing that involves contacting a nucleic acid molecule with a prime editor, a pegRNA, and an inhibitor of the DNA mismatch repair pathway, thereby installing one or more modifications to the nucleic acid molecule at a target site with increased editing efficiency and/or lower indel formation. The present disclosure further provides polynucleotides for editing a DNA target site by prime editing comprising a nucleic acid sequence encoding a napDNAbp, a polymerase, and an inhibitor of the DNA mismatch repair pathway, wherein the napDNAbp and polymerase is capable in the presence of a pegRNA of installing one or more modifications in the DNA target site with increased editing efficiency and/or lower indel formation. Thus, the methods and compositions described herein utilize prime editors, which may comprise a polymerase (e.g., a reverse transcriptase).

[402] In various embodiments, the prime editors used in the methods and compositions disclosed herein includes a polymerase (e.g., DNA-dependent DNA polymerase or RNA-dependent DNA polymerase, such as, reverse transcriptase), or a variant thereof, which can be provided as a fusion protein with a napDNAbp or other programmable nuclease, or provide *in trans*.

[403] Any polymerase may be used in the prime editors with the methods and compositions disclosed herein. The polymerases may be wild type polymerases, functional fragments, mutants, variants, or truncated variants, and the like. The polymerases may include wild type polymerases from eukaryotic, prokaryotic, archael, or viral organisms, and/or the polymerases may be modified by genetic engineering, mutagenesis, directed evolution-based processes. The

polymerases may include T7 DNA polymerase, T5 DNA polymerase, T4 DNA polymerase, Klenow fragment DNA polymerase, DNA polymerase III and the like. The polymerases may also be thermostable, and may include Taq, Tne, Tma, Pfu, Tfl, Tth, Stoffel fragment, VENT® and DEEPVENT® DNA polymerases, KOD, Tgo, JDF3, and mutants, variants and derivatives thereof (see U.S. Pat. No. 5,436,149; U.S. Pat. No. 4,889,818; U.S. Pat. No. 4,965,185; U.S. Pat. No. 5,079,352; U.S. Pat. No. 5,614,365; U.S. Pat. No. 5,374,553; U.S. Pat. No. 5,270,179; U.S. Pat. No. 5,047,342; U.S. Pat. No. 5,512,462; WO 92/06188; WO 92/06200; WO 96/10640; Barnes, W. M., Gene 112:29-35 (1992); Lawyer, F. C., *et al.*, PCR Meth. Appl. 2:275-287 (1993); Flaman, J.-M, *et al.*, Nuc. Acids Res. 22(15):3259-3260 (1994), each of which are incorporated by reference). For synthesis of longer nucleic acid molecules (e.g., nucleic acid molecules longer than about 3-5 Kb in length), at least two DNA polymerases can be employed. In certain embodiments, one of the polymerases can be substantially lacking a 3' exonuclease activity and the other may have a 3' exonuclease activity. Such pairings may include polymerases that are the same or different. Examples of DNA polymerases substantially lacking in 3' exonuclease activity include, but are not limited to, Taq, Tne(exo-), Tma(exo-), Pfu(exo-), Pwo(exo-), exo-KOD and Tth DNA polymerases, and mutants, variants and derivatives thereof.

[404] Preferably, the polymerase usable in the prime editors utilized in the methods and compositions disclosed herein are "template-dependent" polymerase (since the polymerases are intended to rely on the DNA synthesis template to specify the sequence of the DNA strand under synthesis during prime editing. As used herein, the term "template DNA molecule" refers to that strand of a nucleic acid from which a complementary nucleic acid strand is synthesized by a DNA polymerase, for example, in a primer extension reaction of the DNA synthesis template of a PEgRNA.

[405] As used herein, the term "template dependent manner" is intended to refer to a process that involves the template dependent extension of a primer molecule (e.g., DNA synthesis by DNA polymerase). The term "template dependent manner" refers to polynucleotide synthesis of RNA or DNA wherein the sequence of the newly synthesized strand of polynucleotide is dictated by the well-known rules of complementary base pairing (see, for example, Watson, J. D. *et al.*, In: Molecular Biology of the Gene, ^{4}th Ed., W. A. Benjamin, Inc., Menlo Park, Calif. (1987)). The term "complementary" refers to the broad concept of sequence complementarity between regions of two polynucleotide strands or between two nucleotides through base-pairing. It is

known that an adenine nucleotide is capable of forming specific hydrogen bonds ("base pairing") with a nucleotide which is thymine or uracil. Similarly, it is known that a cytosine nucleotide is capable of base pairing with a guanine nucleotide. As such, in the case of prime editing, it can be said that the single strand of DNA synthesized by the polymerase of the prime editor against the DNA synthesis template is said to be "complementary" to the sequence of the DNA synthesis template.

A. Exemplary polymerases

[406] In various embodiments, the prime editors utilized in the methods and compositions described herein comprise a polymerase. The disclosure contemplates any wild type polymerase obtained from any naturally-occurring organism or virus, or obtained from a commercial or non-commercial source. In addition, the polymerases usable in the prime editors can include any naturally-occurring mutant polymerase, engineered mutant polymerase, or other variant polymerase, including truncated variants that retain function. The polymerases usable herein may also be engineered to contain specific amino acid substitutions, such as those specifically disclosed herein. In certain preferred embodiments, the polymerases usable in the prime editors utilized in the methods and compositions of the present disclosure are template-based polymerases, i.e., they synthesize nucleotide sequences in a template-dependent manner.

[407] A polymerase is an enzyme that synthesizes a nucleotide strand and which may be used in connection with the prime editor systems utilized in the methods and compositions described herein. The polymerases are preferably "template-dependent" polymerases (i.e., a polymerase which synthesizes a nucleotide strand based on the order of nucleotide bases of a template strand). In certain configurations, the polymerases can also be a "template-independent" (i.e., a polymerase which synthesizes a nucleotide strand without the requirement of a template strand). A polymerase may also be further categorized as a "DNA polymerase" or an "RNA polymerase." In various embodiments, the prime editor systems comprise a DNA polymerase. In various embodiments, the DNA polymerase can be a "DNA-dependent DNA polymerase" (i.e., whereby the template molecule is a strand of DNA). In such cases, the DNA template molecule can be a PEgRNA, wherein the extension arm comprises a strand of DNA. In such cases, the PEgRNA may be referred to as a chimeric or hybrid PEgRNA which comprises an RNA portion (i.e., the guide RNA components, including the spacer and the gRNA core) and a DNA portion (i.e., the extension arm). In various other embodiments, the DNA polymerase can be an "RNA-

dependent DNA polymerase" (i.e., whereby the template molecule is a strand of RNA). In such cases, the PEgRNA is RNA, i.e., including an RNA extension. The term "polymerase" may also refer to an enzyme that catalyzes the polymerization of nucleotide (i.e., the polymerase activity). Generally, the enzyme will initiate synthesis at the 3'-end of a primer annealed to a polynucleotide template sequence (e.g., such as a primer sequence annealed to the primer binding site of a PEgRNA), and will proceed toward the 5′ end of the template strand. A "DNA polymerase" catalyzes the polymerization of deoxynucleotides. As used herein in reference to a DNA polymerase, the term DNA polymerase includes a "functional fragment thereof". A "functional fragment thereof" refers to any portion of a wild-type or mutant DNA polymerase that encompasses less than the entire amino acid sequence of the polymerase and which retains the ability, under at least one set of conditions, to catalyze the polymerization of a polynucleotide. Such a functional fragment may exist as a separate entity, or it may be a constituent of a larger polypeptide, such as a fusion protein.

[408] In some embodiments, the polymerases can be from bacteriophage. Bacteriophage DNA polymerases are generally devoid of 5′ to 3' exonuclease activity, as this activity is encoded by a separate polypeptide. Examples of suitable DNA polymerases are T4, T7, and phi29 DNA polymerase. The enzymes available commercially are: T4 (available from many sources e.g., Epicentre) and T7 (available from many sources, e.g., Epicentre for unmodified and USB for 3' to 5′ exo T7 "Sequenase" DNA polymerase).

[409] The other embodiments, the polymerases are archaeal polymerases. There are 2 different classes of DNA polymerases which have been identified in archaea: 1. Family B/pol I type (homologs of Pfu from *Pyrococcus furiosus*) and 2. pol II type (homologs of *P. furiosus* DP1/DP2 2-subunit polymerase). DNA polymerases from both classes have been shown to naturally lack an associated 5′ to 3' exonuclease activity and to possess 3' to 5′ exonuclease (proofreading) activity. Suitable DNA polymerases (pol I or pol II) can be derived from archaea with optimal growth temperatures that are similar to the desired assay temperatures. Thermostable archaeal DNA polymerases are isolated from *Pyrococcus* species (furiosus, species GB-D, w*oesii, abysii, horikoshii*), *Thermococcus* species (kodakaraensis KOD1, litoralis, species 9 degrees North-7, species JDF-3, gorgonarius), *Pyrodictium occultum*, and *Archaeoglobus fulgidus*.

[410] Polymerases may also be from eubacterial species. There are 3 classes of eubacterial DNA polymerases, pol I, II, and III. Enzymes in the Pol I DNA polymerase family possess 5′ to 3' exonuclease activity, and certain members also exhibit 3' to 5′ exonuclease activity. Pol II DNA polymerases naturally lack 5′ to 3' exonuclease activity, but do exhibit 3' to 5′ exonuclease activity. Pol III DNA polymerases represent the major replicative DNA polymerase of the cell and are composed of multiple subunits. The pol III catalytic subunit lacks 5′ to '3' exonuclease activity, but in some cases'3' to 5′ exonuclease activity is located in the same polypeptide. There are a variety of commercially available Pol I DNA polymerases, some of which have been modified to reduce or abolish 5′ to'3' exonuclease activity.

[411] Suitable thermostable pol I DNA polymerases can be isolated from a variety of thermophilic eubacteria, including *Thermus* species and *Thermotoga maritima* such as *Thermus aquaticus* (Taq), *Thermus thermophilus* (Tth) and *Thermotoga maritima* (Tma UlTma). Additional eubacteria related to those listed above are described in Thermophilic Bacteria (Kristjansson, J. K., ed.) CRC Press, Inc., Boca Raton, Fla., 1992.

[412] The invention further provides for chimeric or non-chimeric DNA polymerases that are chemically modified according to methods disclosed in U.S. Pat. Nos. 5,677,152, 6,479,264 and 6,183,998, the contents of which are hereby incorporated by reference in their entirety. Additional archaea DNA polymerases related to those listed above are described in the following references: Archaea: A Laboratory Manual (Robb, F. T. and Place, A. R., eds.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1995 and Thermophilic Bacteria (Kristjansson, J. K., ed.) CRC Press, Inc., Boca Raton, Fla., 1992.

B. Exemplary reverse transcriptases

[413] In various embodiments, the prime editors utilized in the methods and compositions described herein comprise a reverse transcriptase as the polymerase. The disclosure contemplates any wild type reverse transcriptase obtained from any naturally-occurring organism or virus, or obtained from a commercial or non-commercial source. In addition, the reverse transcriptases usable in the prime editors utilized in the methods and compositions of the disclosure can include any naturally-occurring mutant RT, engineered mutant RT, or other variant RT, including truncated variants that retain function. The RTs may also be engineered to contain specific amino acid substitutions, such as those specifically disclosed herein.

[414] Reverse transcriptases are multi-functional enzymes typically with three enzymatic activities including RNA- and DNA-dependent DNA polymerization activity, and an RNaseH activity that catalyzes the cleavage of RNA in RNA-DNA hybrids. Some mutants of reverse transcriptases have disabled the RNaseH moiety to prevent unintended damage to the mRNA. These enzymes that synthesize complementary DNA (cDNA) using mRNA as a template were first identified in RNA viruses. Subsequently, reverse transcriptases were isolated and purified directly from virus particles, cells or tissues. (e.g., see Kacian *et al.*, 1971, Biochim. Biophys. Acta 46: 365-83; Yang *et al.*, 1972, Biochem. Biophys. Res. Comm. 47: 505-11; Gerard *et al.*, 1975, J. Virol. 15: 785-97; Liu *et al.*, 1977, Arch. Virol. 55 187-200; Kato *et al.*, 1984, J. Virol. Methods 9: 325-39; Luke *et al.*, 1990, Biochem. 29: 1764-69 and Le Grice *et al.*, 1991, J. Virol. 65: 7004-07, each of which are incorporated by reference). More recently, mutants and fusion proteins have been created in the quest for improved properties such as thermostability, fidelity and activity. Any of the wild type, variant, and/or mutant forms of reverse transcriptase which are known in the art or which can be made using methods known in the art are contemplated herein.

[415] The reverse transcriptase (RT) gene (or the genetic information contained therein) can be obtained from a number of different sources. For instance, the gene may be obtained from eukaryotic cells which are infected with retrovirus, or from a number of plasmids which contain either a portion of or the entire retrovirus genome. In addition, messenger RNA-like RNA which contains the RT gene can be obtained from retroviruses. Examples of sources for RT include, but are not limited to, Moloney murine leukemia virus (M-MLV or MLVRT); human T-cell leukemia virus type 1 (HTLV-1); bovine leukemia virus (BLV); Rous Sarcoma Virus (RSV); human immunodeficiency virus (HIV); yeast, including Saccharomyces, Neurospora, Drosophila; primates; and rodents. See, for example, Weiss, *et al.*, U.S. Pat. No. 4,663,290 (1987); Gerard, G. R., DNA:271-79 (1986); Kotewicz, M. L., *et al.*, Gene 35:249-58 (1985); Tanese, N., *et al.*, Proc. Natl. Acad. Sci. (USA):4944-48 (1985); Roth, M. J., at al., J. Biol. Chem. 260:9326-35 (1985); Michel, F., *et al.*, Nature 316:641-43 (1985); Akins, R. A., *et al.*, Cell 47:505-16 (1986), EMBO J. 4:1267-75 (1985); and Fawcett, D. F., Cell 47:1007-15 (1986) (each of which are incorporated herein by reference in their entireties).

Wild type RTs

[416] Exemplary enzymes for use with the prime editors can include, but are not limited to, M-MLV reverse transcriptase and RSV reverse transcriptase. Enzymes having reverse transcriptase activity are commercially available. In certain embodiments, the reverse transcriptase provided *in trans* to the other components of the prime editor system. That is, the reverse transcriptase is expressed or otherwise provided as an individual component, i.e., not as a fusion protein with a napDNAbp.

[417] A person of ordinary skill in the art will recognize that wild type reverse transcriptases, including but not limited to, Moloney Murine Leukemia Virus (M-MLV); Human Immunodeficiency Virus (HIV) reverse transcriptase and avian Sarcoma-Leukosis Virus (ASLV) reverse transcriptase, which includes but is not limited to Rous Sarcoma Virus (RSV) reverse transcriptase, Avian Myeloblastosis Virus (AMV) reverse transcriptase, Avian Erythroblastosis Virus (AEV) Helper Virus MCAV reverse transcriptase, Avian Myelocytomatosis Virus MC29 Helper Virus MCAV reverse transcriptase, Avian Reticuloendotheliosis Virus (REV-T) Helper Virus REV-A reverse transcriptase, Avian Sarcoma Virus UR2 Helper Virus UR2AV reverse transcriptase, Avian Sarcoma Virus Y73 Helper Virus YAV reverse transcriptase, Rous Associated Virus (RAV) reverse transcriptase, and Myeloblastosis Associated Virus (MAV) reverse transcriptase may be suitably used in the subject methods and composition described herein.

[418] Exemplary wild type RT enzymes are as follows:

Description	Sequence
REVERSE TRANSCRIPTASE (M-MLV RT) WILD TYPE MOLONEY MURINE LEUKEMIA VIRUS USED IN PE1 (PRIME EDITOR 1 FUSION PROTEIN DISCLOSED HEREIN)	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATST PVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQD LREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFE WRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLL LAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLT EARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNW GPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRP VAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVK QPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILA EAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAG TSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEI KNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDT STLLIENSSP(SEQ ID NO: 81)
REVERSE TRANSCRIPTASE	AFPLERPDWDYTTQAGRNHLVHYRQLLLAGLQNAGRSPTNLAKVKGITQGPNESP SAFLERLKEAYRRYTPYDPEDPGQETNVSMSFIWQSAPDIGRKLGRLEDLKSKTLG DLVREAEKIFNKRETPEEREERIRRETEEKEERRRTVDEQKEKERDRRRHREMSKLL ATVVIGQEQDRQEGERKRPQLDKDQCAYCKEKGHWAKDCPKKPRGPRGPRPQTS

Description	Sequence
MOLONEY MURINE LEUKEMIA VIRUS REF SEQ. AAA66622.1	LLTLGDXGGQGQDPPPEPRITLKVGGQPVTFLVDTGAQHSVLTQNPGPLSDKSAW VQGATGGKRYRWTTDRKVHLATGKVTHSFLHVPDCPYPLLGRDLLTKLKAQIHFE GSGAQVVGPMGQPLQVLTLNIEDEYRLHETSKEPDVSLGFTWLSDFPQAWAESGG MGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNT PLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLK DAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLA DFR(SEQ ID NO: 69)
REVERSE TRANSCRIPTASE FELINE LEUKEMIA VIRUS REF SEQ. NP955579.1	TLQLEEEYRLFEPESTQKQEMDIWLKNFPQAWAETGGMGTAHCQAPVLIQLKATA TPISIRQYPMPHEAYQGIKPHIRRMLDQGILKPCQSPWNTPLLPVKKPGTEDYRPVQ DLREVNKRVEDIHPTVPNPYNLLSTLPPSHPWYTVLDLKDAFFCLRLHSESQLLFAF EWRDPEIGLSGQLTWTRLPQGFKNSPTLFDEALHSDLADFRVRYPALVLLQYVDDL LLAAATRTECLEGTKALLETLGNKGYRASAKKAQICLQEVTYLGYSLKDGQRWLT KARKEAILSIPVPKNSRQVREFLGTAGYCRLWIPGFAELAAPLYPLTRPGTLFQWGT EQQLAFEDIKKALLSSPALGLPDITKPFELFIDENSGFAKGVLVQKLGPWKRPVAYL SKKLDTVASGWPPCLRMVAAIAILVKDAGKLTLGQPLTILTSHPVEALVRQPPNKW LSNARMTHYQAMLLDAERVHFGPTVSLNPATLLPLPSGGNHHDCLQILAETHGTRP DLTDQPLPDADLTWYTDGSSFIRNGEREAGAAVTTESEVIWAAPLPPGTSAQRAELI ALTQALKMAEGKKLTVYTDSRYAFATTHVHGEIYRRRGLLTSEGKEIKNKNEILAL LEALFLPKRLSIIHCPGHQKGDSPQAKGNRLADDTAKKAATETHSSLTVL SEQ ID NO: 70)
REVERSE TRANSCRIPTASE HIV-1 RT, CHAIN A REF SEQ. ITL3-A	PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNT PVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGD AYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPF RKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPF LWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLXK LLRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWT YQIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQ KETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETFYVDGAANR ETKLGKAGYVTNRGRQKVVTLTDTTNQKTELQAIYLALQDSGLEVNIVTDSQYAL GIIQAQPDQSESELVNQIIEQLIKKEKVYLAWVPAHKGIGGNEQVDKLVSAGIRKV(SEQ ID NO: 71) See Martinelli *et al.*, Virology, 1990, 174(1): 135-144, which is incorporated by reference
REVERSE TRANSCRIPTASE HIV-1 RT, CHAIN B REF SEQ. ITL3-B	PISPIETVPVKLKPGMDGPKVKQWPLTEEKIKALVEICTEMEKEGKISKIGPENPYNT PVFAIKKKDSTKWRKLVDFRELNKRTQDFWEVQLGIPHPAGLKKKKSVTVLDVGD AYFSVPLDEDFRKYTAFTIPSINNETPGIRYQYNVLPQGWKGSPAIFQSSMTKILEPF RKQNPDIVIYQYMDDLYVGSDLEIGQHRTKIEELRQHLLRWGLTTPDKKHQKEPPF LWMGYELHPDKWTVQPIVLPEKDSWTVNDIQKLVGKLNWASQIYPGIKVRQLCKL LRGTKALTEVIPLTEEAELELAENREILKEPVHGVYYDPSKDLIAEIQKQGQGQWTY QIYQEPFKNLKTGKYARMRGAHTNDVKQLTEAVQKITTESIVIWGKTPKFKLPIQK ETWETWWTEYWQATWIPEWEFVNTPPLVKLWYQLEKEPIVGAETF(SEQ ID NO: 72) See Stammers *et al.*, J. Mol. Biol., 1994, 242(4): 586-588, which is incorporated by reference
REVERSE TRANSCRIPTASE ROUS SARCOMA VIRUS RT REF SEQ. ACL14945	TVALHLAIPLKWKPNHTPVWIDQWPLPEGKLVALTQLVEKELQLGHIEPSLSCWNT PVFVIRKASGSYRLLHDLRAVNAKLVPFGAVQQGAPVLSALPRGWPLMVLDLKDC FFSIPLAEQDREAFAFTLPSVNNQAPARRFQWKVLPQGMTCSPTICQLIVGQILEPLR LKHPSLRMLHYMDDLLLAASSHDGLEAAGEEVISTLERAGFTISPDKVQKEPGVQY LGYKLGSTYAAPVGLVAEPRIATLWDVQKLVGSLQWLRPALGIPPRLRGPFYEQLR GSDPNEAREWNLDMKMAWREIVQLSTTAALERWDPALPLEGAVARCEQGAIGVL GQGLSTHPRPCLWLFSTQPTKAFTAWLEVLTLLITKLRASAVRTFGKEVDILLLPAC FRDELPLPEGILLALRGFAGKIRSSDTPSIFDIARPLHVSLKVRVTDHPVPGPTVFTDA SSSTHKGVVVWREGPRWEIKEIADLGASVQQLEARAVAMALLLWPTTPTNVVTDS

Description	Sequence
	AFVAKMLLKMGQEGVPSTAAAFILEDALSQRSAMAAVLHVRSHSEVPGFFTEGND VADSQATFQAYPLREAKDLHTALHIGPRALSKACNISMQQARFVVQTCPHCNSAP ALEAGVNPRGLGPLQIWQTDFTLEPRMAPRSWLAVTVDTASSAIVVTQHGRVTSV AAQHHWATVIAVLGRPKAIKTDNGSCFTSKSTREWLARWGIAHTTGIPGNSQGQA MVERANRLLKDKIRVLAEGDGFMKRIPTSKQGELLAKAMYALNHFERGENTKTPI QKHWRPTVLTEGPPVKIRIETGEWEKGWNVLVWGRGYAAVKNRDTDKVIWVPSR KVKPDIAQKDEVTKKDEASPLFA(SEQ ID NO: 73) See Yasukawa *et al.*, J. Biochem. 2009, 145(3): 315-324, which is incorporated by reference
REVERSE TRANSCRIPTASE CAULIFLOWER MOSAIC VIRUS RT REF SEQ. AGT42196	MMDHLLQKTQIQNQTEQVMNITNPNSIYIKGRLYFKGYKKIELHCFVDTGASLCIA SKFVIPEEHWINAERPIMVKIADGSSITINKVCRDIDLIIAGEIFHIPTVYQQESGIDFII GNNFCQLYEPFIQFTDRVIFTKDRTYPVHIAKLTRAVRVGTEGFLESMKKRSKTQQP EPVNISTNKIAILSEGRRLSEEKLFITQQRMQKIEELLEKVCSENPLDPNKTKQWMK ASIKLSDPSKAIKVKPMKYSPMDREEFDKQIKELLDLKVIKPSKSPHMAPAFLVNNE AEKRRGKKRMVVNYKAMNKATVGDAYNLPNKDELLTLIRGKKIFSSFDCKSGFW QVLLDQDSRPLTAFTCPQGHYEWNVVPFGLKQAPSIFQRHMDEAFRVFRKFCCVY VDDILVFSNNEEDHLLHVAMILQKCNQHGIILSKKKAQLFKKKINFLGLEIDEGTHK PQGHILEHINKFPDTLEDKKQLQRFLGILTYASDYIPKLAQIRKPLQAKLKENVPWK WTKEDTLYMQKVKKNLQGFPPLHHPLPEEKLIIETDASDDYWGGMLKAIKINEGT NTELICRYASGSFKAAEKNYHSNDKETLAVINTIKKFSIYLTPVHFLIRTDNTHFKSF VNLNYKGDSKLGRNIRWQAWLSHYSFDVEHIKGTDNHFADFLSREFNRVNS(SEQ ID NO: 74) See Farzadfar *et al.*, Virus Genes, 2013, 47(2): 347-356, which is incorporated by reference
REVERSE TRANSCRIPTASE KLEBSIELLA PNEUMONIA REF SEQ. RFF81513.1	MKEKISKIDKNFYTDIFIKTSFQNEFEAGGVIPPIAKNQVSTISNKNKTFYSLAHSSPH YSIQTRIEKFLLKNIPLSASSFAFRKERSYLHYLEPHTQNVKYCHLDIVSFFHSIDVNI VRDTFSVYFSDEFLVKEKQSLLDAFMASVTLTAELDGVEKTFIPMGFKSSPSISNIIF RKIDILIQKFCDKNKITYTRYADDLLFSTKKENNILSSTFFINEISSILSINKFKLNKSK YLYKEGTISLGGYVIENILKDNSSGNIRLSSSKLNPLYKALYEIKKGSSSKHICIKVFN LKLKRFIYKKNKEKFEAKFYSSQLKNKLLGYRSYLLSFVIFHKKYKCINPIFLEKCVF LISEIESIMNRKF(SEQ ID NO: 75)
REVERSE TRANSCRIPTASE ESCHERICHIA COLI RT REF SEQ. TGH57013	MKITSNNVTAVINGKGWHSINWKKCHQHVKTIQTRIAKAACNQQWRTVGRLQRL LVRSFSARALAVKRVTENSGRKTPGVDGQIWSTPESKWEAIFKLRRKGYKPLPLKR VFIPKSNGKKRPLGIPVMLDRAMQALHLLGLEPVSETNADHNSYGFRPARCTADAI QQVCNMYSSRNASKWVLEGDIKGCFEHISHEWLLENIPMDKQILRNWLKAGIIEKS IFSKTLSGTPQGGIISPVLANMALDGLERLLQNRFGRNRLI(SEQ ID NO: 76)
REVERSE TRANSCRIPTASE BACILLUS SUBTILIS RT REF SEQ. QBJ66766	MSKIKINYEKYHIKPFPHFDQRIKVNKKVKENLQNPFYIAAHSFYPFIHYKKISYKFK NGTLSSPKERDIFYSGHMDGYIYKHYGEILNHKYNNTCIGKGIDHVSLAYRNNKMG KSNIHFAAEVINFISEQQQAFIFVSDFSSYFDSLDHAILKEKLIEVLEEQDKLSKDWW NVFKHITRYNWVEKEEVISDLECTKEKIARDKKSRERYYTPAEFREFRKRVNIKSND TGVGIPQGTAISAVLANVYAIDLDQKLNQYALKYGGIYRRYSDDIIMVLPMTSDGQ DPSNDHVSFIKSVVKRNKVTMGDSKTSVLYYANNNIYEDYQRKRESKMDYLGFSF DGMTVKIREKSLFKYYHRTYKKINSINWASVKKEKKVGRKKLYLLYSHLGRNYKG HGNFISYCKKAHAVFEGNKKIESLINQQIKRHWKKIQKRLVDV(SEQ ID NO: 77)
EUBACTERIUM RECTALE GROUP II INTRON RT	DTSNLMEQILSSDNLNRAYLQVVRNKGAEGVDGMKYTELKEHLAKNGETIKGQL RTRKYKPQPARRVEIPKPDGGVRNLGVPTVTDRFIQQAIAQVLTPIYEEQFHDHSYG FRPNRCAQQAILTALNIMNDGNDWIVDIDLEKFFDTVNHDKLMTLIGRTIKDGDVIS IVRKYLVSGIMIDDEYEDSIVGTPQGGNLSPLLANIMLNELDKEMEKRGLNFVRYA DDCIIMVGSEMSANRVMRNISRFIEEKLGLKVNMTKSKVDRPSGLKYLGFGFYFDP

Description	Sequence
	RAHQFKAKPHAKSVAKFKKRMKELTCRSWGVSNSYKVEKLNQLIRGWINYFKIGS MKTLCKELDSRIRYRLRMCIWKQWKTPQNQEKNLVKLGIDRNTARRVAYTGKRIA YVCNKGAVNVAISNKRLASFGLISMLDYYIEKCVTC(SEQ ID NO: 78)
GEOBACILLUS STEAROTHERMOPHILUS GROUP II INTRON RT	ALLERILARDNLITALKRVEANQGAPGIDGVSTDQLRDYIRAHWSTIHAQLLAGTY RPAPVRRVEIPKPGGGTRQLGIPTVVDRLIQQAILQELTPIFDPDFSSSSFGFRPGRNA HDAVRQAQGYIQEGYRYVVDMDLEKFFDRVNHDILMSRVARKVKDKRVLKLIRA YLQAGVMIEGVKVQTEEGTPQGGPLSPLLANILLDDLDKELEKRGLKFCRYADDC NIYVKSLRAGQRVKQSIQRFLEKTLKLKVNEEKSAVDRPWKRAFLGFSFTPERKAR IRLAPRSIQRLKQRIRQLTNPNWSISMPERIHRVNQYVMGWIGYFRLVETPSVLQTIE GWIRRRLRLCQWLQWKRVRTRIRELRALGLKETAVMEIANTRKGAWRTTKTPQL HQALGKTYWTAQGLKSLTQR(SEQ ID NO: 79)

Variant and error-prone RTs

[419] Reverse transcriptases are essential for synthesizing complementary DNA (cDNA) strands from RNA templates. Reverse transcriptases are enzymes composed of distinct domains that exhibit different biochemical activities. The enzymes catalyze the synthesis of DNA from an RNA template, as follows: In the presence of an annealed primer, reverse transcriptase binds to an RNA template and initiates the polymerization reaction. RNA-dependent DNA polymerase activity synthesizes the complementary DNA (cDNA) strand, incorporating dNTPs. RNase H activity degrades the RNA template of the DNA:RNA complex. Thus, reverse transcriptases comprise (a) a binding activity that recognizes and binds to a RNA/DNA hybrid, (b) an RNA-dependent DNA polymerase activity, and (c) an RNase H activity. In addition, reverse transcriptases generally are regarded as having various attributes, including their thermostability, processivity (rate of dNTP incorporation), and fidelity (or error-rate). The reverse transcriptase variants contemplated herein may include any mutations to reverse transcriptase that impacts or changes any one or more of these enzymatic activities (e.g., RNA-dependent DNA polymerase activity, RNase H activity, or DNA/RNA hybrid-binding activity) or enzyme properties (e.g., thermostability, processivity, or fidelity). Such variants may be available in the art in the public domain, available commercially, or may be made using known methods of mutagenesis, including directed evolutionary processes (e.g., PACE or PANCE).

[420] In various embodiments, the reverse transcriptase may be a variant reverse transcriptase. As used herein, a "variant reverse transcriptase" includes any naturally occurring or genetically engineered variant comprising one or more mutations (including singular mutations, inversions, deletions, insertions, and rearrangements) relative to a reference sequences (e.g., a reference wild type sequence). An RT may have several activities, including an RNA-dependent DNA polymerase activity, ribonuclease H activity, and DNA-dependent DNA polymerase activity.

Collectively, these activities enable the enzyme to convert single-stranded RNA into double-stranded cDNA. In retroviruses and retrotransposons, this cDNA can then integrate into the host genome, from which new RNA copies can be made via host-cell transcription. Variant RT's may comprise a mutation which impacts one or more of these activities (either which reduces or increases these activities, or which eliminates these activities all together). In addition, variant RTs may comprise one or more mutations which render the RT more or less stable, less prone to aggregation, and facilitates purification and/or detection, and/or other the modification of properties or characteristics.

[421] A person of ordinary skill in the art will recognize that variant reverse transcriptases derived from other reverse transcriptases, including but not limited to Moloney Murine Leukemia Virus (M-MLV); Human Immunodeficiency Virus (HIV) reverse transcriptase and avian Sarcoma-Leukosis Virus (ASLV) reverse transcriptase, which includes but is not limited to Rous Sarcoma Virus (RSV) reverse transcriptase, Avian Myeloblastosis Virus (AMV) reverse transcriptase, Avian Erythroblastosis Virus (AEV) Helper Virus MCAV reverse transcriptase, Avian Myelocytomatosis Virus MC29 Helper Virus MCAV reverse transcriptase, Avian Reticuloendotheliosis Virus (REV-T) Helper Virus REV-A reverse transcriptase, Avian Sarcoma Virus UR2 Helper Virus UR2AV reverse transcriptase, Avian Sarcoma Virus Y73 Helper Virus YAV reverse transcriptase, Rous Associated Virus (RAV) reverse transcriptase, and Myeloblastosis Associated Virus (MAV) reverse transcriptase may be suitably used in the subject methods and composition described herein.

[422] One method of preparing variant RTs is by genetic modification (e.g., by modifying the DNA sequence of a wild-type reverse transcriptase). A number of methods are known in the art that permit the random as well as targeted mutation of DNA sequences (see for example, Ausubel et. al. *Short Protocols in Molecular Biology* (1995) 3.sup.rd Ed. John Wiley & Sons, Inc.). In addition, there are a number of commercially available kits for site-directed mutagenesis, including both conventional and PCR-based methods. Examples include the QuikChange Site-Directed Mutagenesis Kits (AGILENT®), the Q5® Site-Directed Mutagenesis Kit (NEW ENGLAND BIOLABS®), and GeneArt™ Site-Directed Mutagenesis System (THERMOFISHER SCIENTIFIC®).

[423] In addition, mutant reverse transcriptases may be generated by insertional mutation or truncation (N-terminal, internal, or C-terminal insertions or truncations) according to

methodologies known to one skilled in the art. The term "mutation," as used herein, refers to a substitution of a residue within a sequence, e.g., a nucleic acid or amino acid sequence, with another residue, or a deletion or insertion of one or more residues within a sequence. Mutations are typically described herein by identifying the original residue followed by the position of the residue within the sequence and by the identity of the newly substituted residue. Various methods for making the amino acid substitutions (mutations) provided herein are well known in the art, and are provided by, for example, Green and Sambrook, Molecular Cloning: A Laboratory Manual (4th ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)). Mutations can include a variety of categories, such as single base polymorphisms, microduplication regions, indel, and inversions, and is not meant to be limiting in any way. Mutations can include "loss-of-function" mutations which is the normal result of a mutation that reduces or abolishes a protein activity. Most loss-of-function mutations are recessive, because in a heterozygote the second chromosome copy carries an unmutated version of the gene coding for a fully functional protein whose presence compensates for the effect of the mutation. Mutations also embrace "gain-of-function" mutations, which is one which confers an abnormal activity on a protein or cell that is otherwise not present in a normal condition. Many gain-of-function mutations are in regulatory sequences rather than in coding regions, and can therefore have a number of consequences. For example, a mutation might lead to one or more genes being expressed in the wrong tissues, these tissues gaining functions that they normally lack. Because of their nature, gain-of-function mutations are usually dominant.

[424] Older methods of site-directed mutagenesis known in the art rely on sub-cloning of the sequence to be mutated into a vector, such as an M13 bacteriophage vector, that allows the isolation of single-stranded DNA template. In these methods, one anneals a mutagenic primer (i.e., a primer capable of annealing to the site to be mutated but bearing one or more mismatched nucleotides at the site to be mutated) to the single-stranded template and then polymerizes the complement of the template starting from the 3′ end of the mutagenic primer. The resulting duplexes are then transformed into host bacteria and plaques are screened for the desired mutation.

[425] More recently, site-directed mutagenesis has employed PCR methodologies, which have the advantage of not requiring a single-stranded template. In addition, methods have been developed that do not require sub-cloning. Several issues must be considered when PCR-based

site-directed mutagenesis is performed. First, in these methods it is desirable to reduce the number of PCR cycles to prevent expansion of undesired mutations introduced by the polymerase. Second, a selection must be employed in order to reduce the number of non-mutated parental molecules persisting in the reaction. Third, an extended-length PCR method is preferred in order to allow the use of a single PCR primer set. And fourth, because of the non-template-dependent terminal extension activity of some thermostable polymerases it is often necessary to incorporate an end-polishing step into the procedure prior to blunt-end ligation of the PCR-generated mutant product.

[426] Methods of random mutagenesis, which will result in a panel of mutants bearing one or more randomly situated mutations, exist in the art. Such a panel of mutants may then be screened for those exhibiting the desired properties, for example, increased stability, relative to a wild-type reverse transcriptase.

[427] An example of a method for random mutagenesis is the so-called "error-prone PCR method." As the name implies, the method amplifies a given sequence under conditions in which the DNA polymerase does not support high fidelity incorporation. Although the conditions encouraging error-prone incorporation for different DNA polymerases vary, one skilled in the art may determine such conditions for a given enzyme. A key variable for many DNA polymerases in the fidelity of amplification is, for example, the type and concentration of divalent metal ion in the buffer. The use of manganese ion and/or variation of the magnesium or manganese ion concentration may therefore be applied to influence the error rate of the polymerase.

[428] In various aspects, the RT of the prime editors may be an "error-prone" reverse transcriptase variant. Error-prone reverse transcriptases that are known and/or available in the art may be used. The error-rate of any particular reverse transcriptase is a property of the enzyme's "fidelity," which represents the accuracy of template-directed polymerization of DNA against its RNA template. An RT with high fidelity has a low-error rate. Conversely, an RT with low fidelity has a high-error rate. The fidelity of M-MLV-based reverse transcriptases are reported to have an error rate in the range of one error in 15,000 to 27,000 nucleotides synthesized. See Boutabout *et al.*, "DNA synthesis fidelity by the reverse transcriptase of the yeast retrotransposon Ty1," *Nucleic Acids Res*, 2001, 29: 2217-2222, which is incorporated by reference. Thus, for purposes of this application, those reverse transcriptases considered to be

"error-prone" or which are considered to have an "error-prone fidelity" are those having an error rate that is less than one error in 15,000 nucleotides synthesized.

Error-prone reverse transcriptase also may be created through mutagenesis of a starting RT enzyme (e.g., a wild type M-MLV RT). The method of mutagenesis is not limited and may include directed evolution processes, such as phage-assisted continuous evolution (PACE) or phage-assisted noncontinuous evolution (PANCE). The term "phage-assisted continuous evolution (PACE)," as used herein, refers to continuous evolution that employs phage as viral vectors. The general concept of PACE technology has been described, for example, in International PCT Application, PCT/US2009/056194, filed September 8, 2009, published as WO 2010/028347 on March 11, 2010; International PCT Application, PCT/US2011/066747, filed December 22, 2011, published as WO 2012/088381 on June 28, 2012; U.S. Application, U.S. Patent No. 9,023,594, issued May 5, 2015, International PCT Application, PCT/US2015/012022, filed January 20, 2015, published as WO 2015/134121 on September 11, 2015, and International PCT Application, PCT/US2016/027795, filed April 15, 2016, published as WO 2016/168631 on October 20, 2016, the entire contents of each of which are incorporated herein by reference.

[429] Error-prone reverse transcriptases may also be obtain by phage-assisted non-continuous evolution (PANCE)," which as used herein, refers to non-continuous evolution that employs phage as viral vectors. PANCE is a simplified technique for rapid in vivo directed evolution using serial flask transfers of evolving 'selection phage' (SP), which contain a gene of interest to be evolved, across fresh *E. coli* host cells, thereby allowing genes inside the host *E. coli* to be held constant while genes contained in the SP continuously evolve. Serial flask transfers have long served as a widely-accessible approach for laboratory evolution of microbes, and, more recently, analogous approaches have been developed for bacteriophage evolution. The PANCE system features lower stringency than the PACE system.

[430] Other error-prone reverse transcriptases have been described in the literature, each of which are contemplated for use in the herein methods and compositions. For example, error-prone reverse transcriptases have been described in Bebenek *et al.*, "Error-prone Polymerization by HIV-1 Reverse Transcriptase," *J Biol Chem*, 1993, Vol. 268: 10324-10334 and Sebastian-Martin *et al.*, "Transcriptional inaccuracy threshold attenuates differences in RNA-dependent DNA synthesis fidelity between retroviral reverse transcriptases," *Scientific Reports*, 2018, Vol. 8: 627, each of which are incorporated by reference. Still further, reverse transcriptases,

including error-prone reverse transcriptases can be obtained from a commercial supplier, including ProtoScript® (II) Reverse Transcriptase, AMV Reverse Transcriptase, WarmStart® Reverse Transcriptase, and M-MuLV Reverse Transcriptase, all from NEW ENGLAND BIOLABS®, or AMV Reverse Transcriptase XL, SMARTScribe Reverse Transcriptase, GPR ultra-pure MMLV Reverse Transcriptase, all from TAKARA BIO USA, INC. (formerly CLONTECH).

[431] The herein disclosure also contemplates reverse transcriptases having mutations in RNaseH domain. As mentioned above, one of the intrinsic properties of reverse transcriptases is the RNase H activity, which cleaves the RNA template of the RNA:cDNA hybrid concurrently with polymerization. The RNase H activity can be undesirable for synthesis of long cDNAs because the RNA template may be degraded before completion of full-length reverse transcription. The RNase H activity may also lower reverse transcription efficiency, presumably due to its competition with the polymerase activity of the enzyme. Thus, the present disclosure contemplates any reverse transcriptase variants that comprise a modified RNaseH activity. The herein disclosure also contemplates reverse transcriptases having mutations in the RNA-dependent DNA polymerase domain. As mentioned above, one of the intrinsic properties of reverse transcriptases is the RNA-dependent DNA polymerase activity, which incorporates the nucleobases into the nascent cDNA strand as coded by the template RNA strand of the RNA:cDNA hybrid. The RNA-dependent DNA polymerase activity can be increased or decreased (i.e., in terms of its rate of incorporation) to either increase or decrease the processivity of the enzyme. Thus, the present disclosure contemplates any reverse transcriptase variants that comprise a modified RNA-dependent DNA polymerase activity such that the processivity of the enzyme of either increased or decreased relative to an unmodified version.

[432] Also contemplated herein are reverse transcriptase variants that have altered thermostability characteristics. The ability of a reverse transcriptase to withstand high temperatures is an important aspect of cDNA synthesis. Elevated reaction temperatures help denature RNA with strong secondary structures and/or high GC content, allowing reverse transcriptases to read through the sequence. As a result, reverse transcription at higher temperatures enables full-length cDNA synthesis and higher yields, which can lead to an improved generation of the 3′ flap ssDNA as a result of the prime editing process. Wild type M-MLV reverse transcriptase typically has an optimal temperature in the range of 37-48°C;

however, mutations may be introduced that allow for the reverse transcription activity at higher temperatures of over 48°C, including 49°C, 50°C, 51°C, 52°C, 53°C, 54°C, 55°C, 56°C, 57°C, 58°C, 59°C, 60°C, 61°C, 62°C, 63°C,64°C,65°C,66°C, and higher.

[433] The variant reverse transcriptases contemplated herein, including error-prone RTs, thermostable RTs, increase-processivity RTs, can be engineered by various routine strategies, including mutagenesis or evolutionary processes. In some cases, the variants can be produced by introducing a single mutation. In other cases, the variants may require more than one mutation. For those mutants comprising more than one mutation, the effect of a given mutation may be evaluated by introduction of the identified mutation to the wild-type gene by site-directed mutagenesis in isolation from the other mutations borne by the particular mutant. Screening assays of the single mutant thus produced will then allow the determination of the effect of that mutation alone.

[434] Variant RT enzymes used herein may also include other "RT variants" having at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to any reference RT protein, including any wild type RT, or mutant RT, or fragment RT, or other variant of RT disclosed or contemplated herein or known in the art.

[435] In some embodiments, an RT variant may have 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 21, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, or up to 100, or up to 200, or up to 300, or up to 400, or up to 500 or more amino acid changes compared to a reference RT. In some embodiments, the RT variant comprises a fragment of a reference RT, such that the fragment is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to the corresponding fragment of the reference RT. In some embodiments, the fragment is at least 30%, at least 35%, at least 40%, at least 45%, at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95% identical, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% of the amino acid length of a

corresponding wild type RT (M-MLV reverse transcriptase) (*e.g.*, SEQ ID NO: 81) or to any of the reverse transcriptases of SEQ ID NOs: 69-79.

[436] In some embodiments, the disclosure also may utilize RT fragments which retain their functionality and which are fragments of any herein disclosed RT proteins. In some embodiments, the RT fragment is at least 100 amino acids in length. In some embodiments, the fragment is at least 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, or up to 600 or more amino acids in length.

[437] In still other embodiments, the disclosure also may utilize RT variants which are truncated at the N-terminus or the C-terminus, or both, by a certain number of amino acids which results in a truncated variant which still retains sufficient polymerase function. In some embodiments, the RT truncated variant has a truncation of at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, or 250 amino acids at the N-terminal end of the protein. In other embodiments, the RT truncated variant has a truncation of at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, or 250 amino acids at the C-terminal end of the protein. In still other embodiments, the RT truncated variant has a truncation at the N-terminal and the C-terminal end which are the same or different lengths. For example, the prime editors utilized in the methods and compositions disclosed herein may include a truncated version of M-MLV reverse transcriptase. In this embodiment, the reverse transcriptase contains 4 mutations (D200N, T306K, W313F, T330P; noting that the L603W mutation present in PE2 is no longer present due to the truncation). The DNA sequence encoding this truncated editor is 522 bp smaller than PE2, and therefore makes its potentially useful for applications where delivery of the DNA sequence is challenging due to its size (i.e., adeno-associated virus and lentivirus delivery). This embodiment is referred to as MMLV-RT(trunc) and has the following amino acid sequence:

MMLV-RT(TRUNC)	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGT NDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCL RLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADF RIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQIC QKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLF IPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKP FELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAA IAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLD TDRVQFGPVVALNPATLLPLPEEGLQHNCLDNSRLIN (SEQ ID NO: 80)

[438] In various embodiments, the prime editors utilized in the methods and compositions disclosed herein may comprise one of the RT variants described herein, or a RT variant thereof having at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to any reference Cas9 variants.

In still other embodiments, the present methods and compositions may utilize a DNA polymerase that has been evolved into a reverse transcriptase, as described in Effefson *et al.*, "Synthetic evolutionary origin of a proofreading reverse transcriptase," *Science*, June 24, 2016, Vol. 352: 1590-1593, the contents of which are incorporated herein by reference.

[439] In certain other embodiments, the reverse transcriptase is provided as a component of a fusion protein also comprising a napDNAbp. In other words, in some embodiments, the reverse transcriptase is fused to a napDNAbp as a fusion protein.

[440] In various embodiments, variant reverse transcriptases can be engineered from wild type M-MLV reverse transcriptase as represented by SEQ ID NO: 81.

[441] In various embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising one or more of the following mutations: P51L, S67K, E69K, L139P, T197A, D200N, H204R, F209N, E302K, E302R, T306K, F309N, W313F, T330P, L345G, L435G, N454K, D524G, E562Q, D583N, H594Q, L603W, E607K, or D653N in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence.

[442] Some exemplary reverse transcriptases that can be fused to napDNAbp proteins or provided as individual proteins according to various embodiments of this disclosure are provided below. Exemplary reverse transcriptases include variants with at least 80%, at least 85%, at least

90%, at least 95%, or at least 99% sequence identity to the following wild-type enzymes or partial enzymes:

Description	Sequence (variant substitutions relative to wild type)
Reverse transcriptase (M-MLV RT) wild type Moloney murine leukemia virus Used in PE1 (prime editor 1 fusion protein disclosed herein)	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 81)
M-MLV RT D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 82)
M-MLV RT D200N T330P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 83)
M-MLV RT D200N T330P L603W	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLF**N**EALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTK**P**GTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRG**W**LTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 84)
M-MLV RT D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQ**K**ARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK

Description	Sequence (variant substitutions relative to wild type)
T330P L603W E69K	RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 85)
M-MLV RT D200N T330P L603W E302R	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLRRFLGTAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP(SEQ ID NO: 86)
M-MLV RT D200N T330P L603W E607K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSKGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGN RMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 87)
M-MLV RT D200N T330P L603W L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGPPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 88)
M-MLV RT D200N T330P L603W L435G	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVIGAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF

Description	Sequence (variant substitutions relative to wild type)
	GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 89)
M-MLV RT D200N T330P L603W N454K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSKARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 90)
M-MLV RT D200N T330P L603W T306K	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGKAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 91)
M-MLV RT D200N T330P L603W W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 92)
M-MLV RT D200N T330P L603W D524G E562Q D583N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTGGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAQLIALTQALKMAEGKKLNVYTNSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 93)
M-MLV RT D200N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK

Description	Sequence (variant substitutions relative to wild type)
T330P L603W E302R W313F	RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLRRFLGTAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 94)
M-MLV RT D200N T330P L603W E607K L139P	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGPPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGTAGFCRLWIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSKGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGN RMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 95)
M-MLV RT P51L S67K T197A H204R E302K F309N W313F T330P L435G N454K D524G D583N H594Q D653N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIILLKATSTPVSI KQYPMKQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPALFDEALRRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQG TRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTP RQLRKFLGTAGNCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVIGAPHAVEALVKQPPDRWLSKARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTGGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTNSRYAFA TAHIQGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MANQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 96)
M-MLV RT D200N P51L S67K T197A H204R E302K F309N W313F T330P L345G N454K D524G D583N H594Q D653N	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIILLKATSTPVSI KQYPMKQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPALFNEALRRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQG TRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTP RQLRKFLGTAGNCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPA LGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVIGAPHAVEALVKQPPDRWLSKARMTHYQALLLDTDRVQF GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTGGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTNSRYAFA TAHIQGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MANQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 97)
M-MLV RT D200N T330P L603W T306K W313F in PE2	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSI KQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNK RVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGIS GQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGT RALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPR QLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPAL GLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVA AIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQF

Description	Sequence (variant substitutions relative to wild type)
	GPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQE GQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNR MADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 98)

[443] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising one or more of the following mutations: P51X, S67X, E69X, L139X, T197X, D200X, H204X, F209X, E302X, T306X, F309X, W313X, T330X, L345X, L435X, N454X, D524X, E562X, D583X, H594X, L603X, E607X, or D653X in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid.

[444] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a P51X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is L.

[445] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a S67X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is K.

[446] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a E69X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is K.

[447] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a L139X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is P.

[448] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a T197X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is A.

[449] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a D200X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is N.

[450] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a H204X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is R.

[451] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a F209X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is N.

[452] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a E302X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is K.

[453] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a E302X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is R.

[454] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a T306X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is K.

[455] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a F309X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is N.

[456] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a W313X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is F.

[457] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a T330X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is P.

[458] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a L345X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is G.

[459] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a L435X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is G.

[460] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a N454X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is K.

[461] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a D524X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is G.

[462] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a E562X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is Q.

[463] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a D583X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is N.

[464] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a H594X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is Q.

[465] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a L603X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is W.

[466] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a E607X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is K.

[467] In various other embodiments, the prime editors utilized in the methods and compositions described herein (with RT provided as either a fusion partner or *in trans*) can include a variant RT comprising a D653X mutation in the wild type M-MLV RT of SEQ ID NO: 81 or at a corresponding amino acid position in another wild type RT polypeptide sequence, wherein "X" can be any amino acid. In certain embodiments, X is N.

[468] Some exemplary reverse transcriptases that can be fused to napDNAbp proteins or provided as individual proteins according to various embodiments of this disclosure are provided below. Exemplary reverse transcriptases include variants with at least 80%, at least 85%, at least 90%, at least 95% or at least 99% sequence identity to the wild-type enzymes or partial enzymes represented by SEQ ID NOs: 81-98.

[469] The prime editor system utilized in the methods and compositions described here contemplates any publicly-available reverse transcriptase described or disclosed in any of the following U.S. patents (each of which are incorporated by reference in their entireties): U.S. Patent Nos: 10,202,658; 10,189,831; 10,150,955; 9,932,567; 9,783,791; 9,580,698; 9,534,201; and 9,458,484, and any variant thereof that can be made using known methods for installing mutations, or known methods for evolving proteins. The following references describe reverse transcriptases in art. Each of their disclosures are incorporated herein by reference in their entireties.

[470] Herzig, E., Voronin, N., Kucherenko, N. & Hizi, A. A Novel Leu92 Mutant of HIV-1 Reverse Transcriptase with a Selective Deficiency in Strand Transfer Causes a Loss of Viral Replication. *J. Virol.* **89**, 8119–8129 (2015).

[471] Mohr, G. *et al.* A Reverse Transcriptase-Cas1 Fusion Protein Contains a Cas6 Domain Required for Both CRISPR RNA Biogenesis and RNA Spacer Acquisition. *Mol. Cell* **72**, 700-714.e8 (2018).

[472] Zhao, C., Liu, F. & Pyle, A. M. An ultraprocessive, accurate reverse transcriptase encoded by a metazoan group II intron. *RNA* **24**, 183–195 (2018).

[473] Zimmerly, S. & Wu, L. An Unexplored Diversity of Reverse Transcriptases in Bacteria. *Microbiol Spectr* **3**, MDNA3-0058–2014 (2015).

[474] Ostertag, E. M. & Kazazian Jr, H. H. Biology of Mammalian L1 Retrotransposons. *Annual Review of Genetics* **35**, 501–538 (2001).

[475] Perach, M. & Hizi, A. Catalytic Features of the Recombinant Reverse Transcriptase of Bovine Leukemia Virus Expressed in Bacteria. *Virology* **259**, 176–189 (1999).

Lim, D. *et al.* Crystal structure of the moloney murine leukemia virus RNase H domain. *J. Virol.* **80**, 8379–8389 (2006).

[476] Zhao, C. & Pyle, A. M. Crystal structures of a group II intron maturase reveal a missing link in spliceosome evolution. *Nature Structural & Molecular Biology* **23**, 558–565 (2016).

Griffiths, D. J. Endogenous retroviruses in the human genome sequence. *Genome Biol.* **2**, REVIEWS1017 (2001).

[477] Baranauskas, A. *et al.* Generation and characterization of new highly thermostable and processive M-MuLV reverse transcriptase variants. *Protein Eng Des Sel* **25**, 657–668 (2012).

[478] Zimmerly, S., Guo, H., Perlman, P. S. & Lambowitz, A. M. Group II intron mobility occurs by target DNA-primed reverse transcription. *Cell* **82**, 545–554 (1995).

[479] Feng, Q., Moran, J. V., Kazazian, H. H. & Boeke, J. D. Human L1 retrotransposon encodes a conserved endonuclease required for retrotransposition. *Cell* **87**, 905–916 (1996).

[480] Berkhout, B., Jebbink, M. & Zsiros, J. Identification of an Active Reverse Transcriptase Enzyme Encoded by a Human Endogenous HERV-K Retrovirus. *Journal of Virology* **73**, 2365–2375 (1999).

[481] Kotewicz, M. L., Sampson, C. M., D'Alessio, J. M. & Gerard, G. F. Isolation of cloned Moloney murine leukemia virus reverse transcriptase lacking ribonuclease H activity. *Nucleic Acids Res* **16**, 265–277 (1988).

[482] Arezi, B. & Hogrefe, H. Novel mutations in Moloney Murine Leukemia Virus reverse transcriptase increase thermostability through tighter binding to template-primer. *Nucleic Acids Res* **37**, 473–481 (2009).

[483] Blain, S. W. & Goff, S. P. Nuclease activities of Moloney murine leukemia virus reverse transcriptase. Mutants with altered substrate specificities. *J. Biol. Chem.* **268**, 23585–23592 (1993).

[484] Xiong, Y. & Eickbush, T. H. Origin and evolution of retroelements based upon their reverse transcriptase sequences. *EMBO J* **9**, 3353–3362 (1990).
Herschhorn, A. & Hizi, A. Retroviral reverse transcriptases. *Cell. Mol. Life Sci.* **67**, 2717–2747 (2010).

[485] Taube, R., Loya, S., Avidan, O., Perach, M. & Hizi, A. Reverse transcriptase of mouse mammary tumour virus: expression in bacteria, purification and biochemical characterization. *Biochem. J.* **329 (Pt 3)**, 579–587 (1998).

[486] Liu, M. *et al.* Reverse Transcriptase-Mediated Tropism Switching in Bordetella Bacteriophage. *Science* **295**, 2091–2094 (2002).

[487] Luan, D. D., Korman, M. H., Jakubczak, J. L. & Eickbush, T. H. Reverse transcription of R2Bm RNA is primed by a nick at the chromosomal target site: a mechanism for non-LTR retrotransposition. *Cell* **72**, 595–605 (1993).

[488] Nottingham, R. M. *et al.* RNA-seq of human reference RNA samples using a thermostable group II intron reverse transcriptase. *RNA* **22**, 597–613 (2016).

[489] Telesnitsky, A. & Goff, S. P. RNase H domain mutations affect the interaction between Moloney murine leukemia virus reverse transcriptase and its primer-template. *Proc. Natl. Acad. Sci. U.S.A.* **90**, 1276–1280 (1993).

[490] Halvas, E. K., Svarovskaia, E. S. & Pathak, V. K. Role of Murine Leukemia Virus Reverse Transcriptase Deoxyribonucleoside Triphosphate-Binding Site in Retroviral Replication and In Vivo Fidelity. *Journal of Virology* **74**, 10349–10358 (2000).

[491] Nowak, E. *et al.* Structural analysis of monomeric retroviral reverse transcriptase in complex with an RNA/DNA hybrid. *Nucleic Acids Res* **41**, 3874–3887 (2013).

[492] Stamos, J. L., Lentzsch, A. M. & Lambowitz, A. M. Structure of a Thermostable Group II Intron Reverse Transcriptase with Template-Primer and Its Functional and Evolutionary Implications. *Molecular Cell* **68**, 926-939.e4 (2017).

[493] Das, D. & Georgiadis, M. M. The Crystal Structure of the Monomeric Reverse Transcriptase from Moloney Murine Leukemia Virus. *Structure* **12**, 819–829 (2004).

[494] Avidan, O., Meer, M. E., Oz, I. & Hizi, A. The processivity and fidelity of DNA synthesis exhibited by the reverse transcriptase of bovine leukemia virus. *European Journal of Biochemistry* **269**, 859–867 (2002).

[495] Gerard, G. F. *et al.* The role of template-primer in protection of reverse transcriptase from thermal inactivation. *Nucleic Acids Res* **30**, 3118–3129 (2002).

[496] Monot, C. *et al.* The Specificity and Flexibility of L1 Reverse Transcription Priming at Imperfect T-Tracts. *PLOS Genetics* **9**, e1003499 (2013).

[497] Mohr, S. *et al.* Thermostable group II intron reverse transcriptase fusion proteins and their use in cDNA synthesis and next-generation RNA sequencing. *RNA* **19**, 958–970 (2013).

[498] Any of the references noted above which relate to reverse transcriptases are hereby incorporated by reference in their entireties, if not already stated so.

Prime editors: fusion proteins

[499] The napDNAbps and the polymerases (e.g., reverse transcriptases) may be provided in the form of a fusion protein. That is, the present disclosure contemplates using prime editors comprising fusion proteins, wherein the fusion proteins comprise a napDNAbp domain and a polymerase (e.g., reverse transcriptase) domain.

[500] The prime editor systems utilized in the methods and compositions described herein contemplate fusion proteins comprising a napDNAbp and a polymerase (e.g., DNA-dependent DNA polymerase or RNA-dependent DNA polymerase, such as, reverse transcriptase), and optionally joined by a linker. The application contemplates any suitable napDNAbp and polymerase (e.g., DNA-dependent DNA polymerase or RNA-dependent DNA polymerase, such as, reverse transcriptase) to be combined in a single fusion protein. Examples of napDNAbps and polymerases (e.g., DNA-dependent DNA polymerase or RNA-dependent DNA polymerase, such as, reverse transcriptase) are each defined herein. Since polymerases are well-known in the art, and the amino acid sequences are readily available, this disclosure is not meant in any way to be limited to those specific polymerases identified herein.

[501] In various embodiments, the fusion proteins may comprise any suitable structural configuration. For example, the fusion protein may comprise from the N-terminus to the C-terminus direction, a napDNAbp fused to a polymerase (e.g., DNA-dependent DNA polymerase or RNA-dependent DNA polymerase, such as, reverse transcriptase) . In other embodiments, the fusion protein may comprise from the N-terminus to the C-terminus direction, a polymerase (e.g., a reverse transcriptase) fused to a napDNAbp. The fused domain may optionally be joined by a linker, e.g., an amino acid sequence. In other embodiments, the fusion proteins may comprise the structure NH_2-[napDNAbp]-[polymerase]-COOH; or NH_2-[polymerase]-

[napDNAbp]-COOH, wherein each instance of "]-[" indicates the presence of an optional linker sequence. In embodiments wherein the polymerase is a reverse transcriptase, the fusion proteins may comprise the structure NH_2-[napDNAbp]-[RT]-COOH, or NH_2-[RT]-[napDNAbp]-COOH, wherein each instance of "]-[" indicates the presence of an optional linker sequence.

[502] In various embodiments, the prime editor fusion protein may have the following amino acid sequence (referred to herein as "PE1"), which includes a Cas9 variant comprising an H840A mutation (i.e., a Cas9 nickase) and an M-MLV RT wild type, as well as an N-terminal NLS sequence (19 amino acids) and an amino acid linker (32 amino acids) that joins the C-terminus of the Cas9 nickase domain to the N-terminus of the RT domain. The PE1 fusion protein has the following structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(wt)]. The amino acid sequence of PE1 and its individual components are as follows:

Description	Sequence
PE1 FUSION PROTEIN CAS9(H840A)-MMLV_RT(WT)	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFK VLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY HLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQ LVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLF GNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVR QQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEK ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDD KVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFM QLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQS FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD ENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNF FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVK KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIK LPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKG SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS*TL NIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPV SIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLR EVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEW RDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPDLILLQYVDDLLL AATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEAR KETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAEMAAPLYPLTKTGTLFNWGPD*

Description	Sequence
	QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYL SKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTR PDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAE LIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILAL LKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPS GGSKRTADGSEFEPKKKRKV (SEQ ID NO: 100) KEY: NUCLEAR LOCALIZATION SEQUENCE (NLS) TOP:(SEQ ID NO: 101), BOTTOM: (SEQ ID NO: 103) **CAS9(H840A)** (SEQ ID NO: 37) **33-AMINO ACID LINKER** (SEQ ID NO: 102) *M-MLV reverse transcriptase* (SEQ ID NO: 81)
PE1 – N-TERMINAL NLS	MKRTADGSEFESPKKKRKV (SEQ ID NO: 101)
PE1 – CAS9 (H840A) (MET MINUS)	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDS GETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEE DKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIK FRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLS KSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTY DDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPI LEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFL KDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGAS AQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFL SGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTY HDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVM KQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDS LTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRH KPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQN EKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSD KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD KAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDF RKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDV RKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIV WDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNP IDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKY VNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADAN LDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTK EVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 37)
PE1 – LINKER BETWEEN CAS9 DOMAIN AND RT DOMAIN (33 AMINO ACIDS)	SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 102)
PE1 – M-MLV RT	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKA TSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTND YRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLH PTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHP DLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVK YLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAE MAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDE KQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDA

Description	Sequence
	GKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPV VALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSL LQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNV YTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCP GHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 81)
PE1 – C-TERMINAL NLS	SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 103)

[503] In another embodiment, the prime editor fusion protein may have the following amino acid sequence (referred to herein as "PE2"), which includes a Cas9 variant comprising an H840A mutation (i.e., a Cas9 nickase) and an M-MLV RT comprising mutations D200N, T330P, L603W, T306K, and W313F, as well as an N-terminal NLS sequence (19 amino acids) and an amino acid linker (33 amino acids) that joins the C-terminus of the Cas9 nickase domain to the N-terminus of the RT domain. The PE2 fusion protein has the following structure: [NLS]-[Cas9(H840A)]-[linker]-[MMLV_RT(D200N)(T330P)(L603W)(T306K)(W313F)]. The amino acid sequence of PE2 is as follows:

PE2 FUSION PROTEIN CAS9(H840A)-MMLV_RT D200N T330P L603W T306K W313F	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKF** **KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL** **QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKY** **PTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD** **KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEK** **KNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIG** **DQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTL** **LKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD** **GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLK** **DNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDK** **GASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEG** **MRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGV** **EDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEE** **RLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFL** **KSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAI** **KKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERM** **KRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDIN** **RLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKN** **YWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKH** **VAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINN** **YHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEI** **GKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRD** **FATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPK** **KYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPI** **DFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELAL** **PSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKR** **VILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTT** **IDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD**<u>**SGGSSGGSSGSETP** **GTSESATPESSGGSSGGSS**</u>*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAET* *GGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSP* *WNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTV* *LDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEAL* *HRDLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKA*

	*QICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFI PGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELF VDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDA GKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVAL NPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQ RKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFA TAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEA RGNRMADQAARKAAITETPDTSTLLIENSSP*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 107) KEY: NUCLEAR LOCALIZATION SEQUENCE (NLS) TOP:(SEQ ID NO: 101), BOTTOM: (SEQ ID NO: 103) **CAS9(H840A)** (SEQ ID NO: 37) **33-AMINO ACID LINKER** (SEQ ID NO: 102) *M-MLV reverse transcriptase* (SEQ ID NO: 98)
PE2 – N-TERMINAL NLS	MKRTADGSEFESPKKKRKV (SEQ ID NO: 101)
PE2 – CAS9 (H840A) (METMINUS)	DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFD SGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLV EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAH MIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSA RLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMI KRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYK FIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQED FYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVV DKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGM RKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRF NASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAH LFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDE LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEH PVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDS IDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKV ITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEF VYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRP LIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKR NSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLG ITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGE LQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKY FDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 37)
PE2 – LINKER BETWEEN CAS9 DOMAIN AND RT DOMAIN (33 AMINO ACIDS)	SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID NO: 102)
PE2 – MMLV_RT D200N T330P L603W T306K W313F	TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLK ATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTN DYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLR LHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRI QHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPG FAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFEL FVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAV

	LTKDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDR VQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTW YTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFL PKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 98)
PE2 – C-TERMINAL NLS	SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 103)

[504] In still other embodiments, the prime editor fusion protein may have the following amino acid sequences:

PE FUSION PROTEIN MMLV_RT(WT)-32AA-CAS9(H840A)	MKRTADGSEFESPKKKRKV*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAET GGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPW NTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDL ADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAE MAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQ PLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLP EEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTE TEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRR GLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARK AAITETPDTSTLLIENSSP**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS**DKKY SIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVE EDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALA HMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAK LQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKA PLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGG ASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGE LHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKS EETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLT LTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEH IANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKG QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMY VDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEV VKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVET RQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKS EQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKG RDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPI DFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPS KYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVIL ADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD*SGGSKRTADGSEFEPKKKR KV (SEQ ID NO: 108) KEY:

	NUCLEAR LOCALIZATION SEQUENCE (NLS) TOP:(SEQ ID NO: 101), BOTTOM: (SEQ ID NO: 103) **CAS9(H840A)** (SEQ ID NO: 37) **33-AMINO ACID LINKER** (SEQ ID NO: 102) *M-MLV reverse transcriptase* (SEQ ID NO: 81)
PE FUSION PROTEIN MMLV_RT(WT)-60AA-CAS9(H840A)	MKRTADGSEFESPKKKRKV*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAET GGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPW NTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDL KDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDL ADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQK QVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAE MAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQG YAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQ PLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLP EEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTE TEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRR GLLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAARK AAITETPDTSTLLIENSSP***SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAG SAAPAAKKKKLDGSGSGGSSGGS*****DKKYSIGLDIGTNSVGWAVITDEYKVPS KKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRIC YLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEK YPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVD KLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKK NGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQ YADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKA LVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEE LLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFI ERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDD KVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFM QLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDE LVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQIL KEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSF LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDEN DKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTAL IKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKT EITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTE VQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVE KGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYS LFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGD*SGGSKRTADGSEFEPKKKRKV* (SEQ ID NO: 109) key: Nuclear localization sequence (NLS) Top:(SEQ ID NO: 101), Bottom: (SEQ ID NO: 103) **Cas9(H840A)**(SEQ ID NO: 37) **amino acid linker** (SEQ ID NO: 128) *M-MLV reverse transcriptase* (SEQ ID NO: 81)
PE FUSION PROTEIN	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFK VLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQE**

CAS9(H840A)-FEN1-MMLV_RT D200N T330P L603W T306K W313F	**IFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIY HLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQ LVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFG NLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLF LAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILT FRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT NFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKK AIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMK QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKV MGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDD SIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIR EVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLA NGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTG GFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKVEKGKS KKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFEL ENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQ LFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAEN IIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRID LSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSGIQGLAKLIADVAP SAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQNEEGETTSHLMGMFY RTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAEKQLQQAQAAGAEQE VEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEASCAALVKAGKVYA AATEDMDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILQELGLNQEQFVDLCIL LGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPVPENWLHKEAHQLFL EPEVLDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGVKRLSKSRQGSTQGR LDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFKRGKSGGSSGGSSGSE TPGTSESATPESSGGSSGGSS*TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAE TGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSP WNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVL DLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHR DLADFRIQHPDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQIC QKQVKYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFA EMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQ GYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMG QPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPL PEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTT ETEVIWAKALPAGTSAQRAELIALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRR RGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARGNRMADQAAR KAAITETPDTSTLLIENSSP*SGGSKRTADGSEFEPKKKRKV (SEQ ID NO: 110) key: Nuclear localization sequence (NLS) Top:(SEQ ID NO: 101), Bottom: (SEQ ID NO: 103) **Cas9(H840A)** (SEQ ID NO: 37) **33-amino acid linker 1** (SEQ ID NO: 102) *M-MLV reverse transcriptase* (SEQ ID NO: 98) **33-amino acid linker 2** (SEQ ID NO: 102) FEN1 (SEQ ID NO: 113)

[505] In other embodiments, the prime editor fusion proteins can be based on SaCas9 or on SpCas9 nickases with altered PAM specificities, such as the following exemplary sequences:

SACAS9-M-MLV RT PRIME EDITOR	MKRTADGSEFESPKKKRKVGKRNYILGLDIGITSVGYGIIDYETRDVIDAGVR LFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGI NPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTKEQI SRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSDYVKEAKQLLKVQKA YHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFP EELRSVKYAYNADLYNALNDLNNLVITRDENEKLEYYEKFQIIENVFKQKKK PTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDITARKEIIENAELL DQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLSLKAINLIL DELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQSI KVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTG KENAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDN SFNNKVLVKQEEASKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISK TKKEYLLEERDINRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVK VKSINGGFTSFLRRKWKFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKA KKVMENQMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDK KPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLM YHHDPQTYQKLKLIMEQYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKI KYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRFDVYLDNGVYKFVTVKNL DVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKINGELYRVIGV NNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNL YEVKSKKHPQIIKKGSGGSSGGSSGSETPGTSESATPESSGGSSGGSSTLNIEDE YRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPV SIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPV QDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTS QPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFDEALHRDLADFRIQHPD LILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVK YLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGTAGFCRLWIPGFAE MAAPLYPLTKTGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFVD EKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTK DAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQ FGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWYT DGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAE GKKLNVYTDSRYAFATAHIHGEIYRRRGLLTSEGKEIKNKDEILALLKALFLP KRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSGG SKRTADGSEFEPKKKRKV (SEQ ID NO: 114)
SPCAS9(H840A)-VRQR-MALONEY MURINE LEUKEMIA VIRUS REVERSE TRANSCRIPTASE PRIME EDITOR	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYN QLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN GIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLH EHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKG QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVD

	QELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVK KMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNY HHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKA TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYN KHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKQYRSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSTLN IEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATS TPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDY RPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLH PTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQH PDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQV KYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFA EMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFV DEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLT KDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRV QFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWY TDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFL PKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSG GSKRTADGSEFEPKKKRKV (SEQ ID NO: 115)
SPCAS9(H840A)-VRER-MALONEY MURINE LEUKEMIA VIRUS REVERSE TRANSCRIPTASE PRIME EDITOR	MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFS NEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRK KLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYN QLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLG LTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADLFLAAKNLSD AILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFF DQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGN SRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKH SLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVK QLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDIL EDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLIN GIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLH EHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKG QKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVD QELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVK KMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITK HVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNY HHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKA TAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRK VLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFVSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEV KKDLIIKLPKYSLFELENGRKRMLASARELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYN KHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKEYRSTKEVLDATLIH QSITGLYETRIDLSQLGGDSGGSSGGSSGSETPGTSESATPESSGGSSGGSSTLN IEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATS TPVSIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDY RPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLH

	PTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQH PDLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQV KYLGYLLKEGQRWLTEARKETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFA EMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFELFV DEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLT KDAGKLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRV QFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTRPDLTDQPLPDADHTWY TDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMA EGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFL PKRLSIIHCPGHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSPSG GSKRTADGSEFEPKKKRKV (SEQ ID NO: 116)

[506] In yet other embodiments, the prime editor fusion proteins utilized in the methods and compositions contemplated herein may include a Cas9 nickase (e.g., Cas9 (H840A)) fused to a truncated version of M-MLV reverse transcriptase. In this embodiment, the reverse transcriptase also contains 4 mutations (D200N, T306K, W313F, T330P; noting that the L603W mutation present in PE2 is no longer present due to the truncation). The DNA sequence encoding this truncated editor is 522 bp smaller than PE2, and therefore makes its potentially useful for applications where delivery of the DNA sequence is challenging due to its size (i.e., adeno-associated virus and lentivirus delivery). This embodiment is referred to as Cas9(H840A)-MMLV-RT(trunc) or "PE2-short"or "PE2-trunc" and has the following amino acid sequence:

CAS9(H840A)-MMLV-RT(TRUNC) OR PE2-SHORT	MKRTADGSEFESPKKKRKV**DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFK VLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQ EIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTI YHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFI QLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGL FGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVR QQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLV KLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEK ILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIE RMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSG EQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGT YHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDD KVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFM QLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVD ELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQI LKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQS FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQ RKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYD ENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNF FKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVK KTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIK LPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKG SPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHR DKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD**SGGSSGGSSGSETPGTSESATPESSGGSSGGSS*TL*

	NIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPV SIKQYPMSQEARLGIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLR EVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDAFFCLRLHPTSQPLFAFEW RDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDLLL AATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEAR KETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPD QQKAYQEIKQALLTAPALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYL SKKLDPVAAGWPPCLRMVAALAVLTKDAGKLTMGQPLVILAPHAVEALVKQPPDR WLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDNSRLINSGG SKRTADGSEFEPKKKRKV (SEQ ID NO: 117) KEY: NUCLEAR LOCALIZATION SEQUENCE (NLS) TOP:(SEQ ID NO: 101), BOTTOM: (SEQ ID NO: 103) **CAS9(H840A)** (SEQ ID NO: 37) **33-AMINO ACID LINKER 1** (SEQ ID NO: 102) *M-MLV TRUNCATED reverse transcriptase* (SEQ ID NO: 80)

[507] In various embodiments, the prime editor fusion proteins utilized in the methods and compositions contemplated herein may also include any variants of the above-disclosed sequences having an amino acid sequence that is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to PE1, PE2, or any of the above indicated prime editor fusion sequences.

[508] In certain embodiments, linkers may be used to link any of the peptides or peptide domains or moieties of the invention (e.g., a napDNAbp linked or fused to a reverse transcriptase).

Prime editors: modified PE fusion proteins (e.g., PEmax)

[509] In one aspect, the present disclosure provides modified prime editor proteins. In one embodiment, the modified prime editor fusion protein is PEmax (of SEQ ID NO: 99), or an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least up to 100% sequence identify with SEQ ID NO: 99. The sequence of PEmax is as follows:

KEY:

Bipartite SV40 NLS

SpCas9 R221K N394K H840A

Linker = (SGGSx2–bipartite SV40 NLS–SGGSx2) (SGGSSGGS KRTADGSEFESPKKKRKV SGGSSGGS) (SEQ ID NO: 105)

Genscript codon optimized MMLV RT pentamutant (D200N T306K W313F T330P L603W)

Other linker sequences

c-Myc NLS

[510] PEmax protein sequence (SEQ ID NO: 99):

MKRTADGSEFESPKKKRKVDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLF
DSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDE
VAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLF
EENPINASGVDAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVR
QQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFDNGSIPH
QIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKG
ASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTN
RKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDR
EMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHD
DSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTT
QKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDA
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYH
HAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLA
NGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLII
KLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYL
DEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTK
EVLDATLIHQSITGLYETRIDLSQLGGDSGGSSGGSKRTADGSEFESPKKKRKVSGGSSGGSTLNIEDEYRLH
ETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARLGIKPHIQRLLD
QGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQWYTVLDLKDA
FFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHPDLILLQYVDDL
LLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEARKETVMGQPTPK
TPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAPALGLPDLTKPFE
LFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAGKLTMGQPLVILA
PHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHNCLDILAEAHGTR
PDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELIALTQALKMAEGK
KLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCPGHQKGHSAEARG
NRMADQAARKAAITETPDTSTLLIENSSPSGGSKRTADGSEFESPKKKRKVGSGPAAKRVKLD

[511] PEmax DNA sequence (SEQ ID NO: 5660):

ATGAAACGGACAGCCGACGGAAGCGAGTTCGAGTCACCAAAGAAGAAGCGGAAAGTCGACAAGAAG
TACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGT
GCCCAGCAAGAAATTCAAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGA
GCCCTGCTGTTCGACAGCGGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGAT
ACACCAGACGGAAGAACCGGATCTGCTATCTGCAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGA
CGACAGCTTCTTCCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAGGATAAGAAGCACGAGCGGCAC
CCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCCACCATCTACCACCTGAG
AAAGAAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCCACATG
ATCAAGTTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCT
GTTCATCCAGCTGGTGCAGACCTACAACCAGCTGTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGG
ACGCCAAGGCCATCCTGTCTGCCAGACTGAGCAAGAGCAGAAAGCTGGAAAATCTGATCGCCCAGCT

GCCCGGCGAGAAGAAGAATGGCCTGTTCGGAAACCTGATTGCCCTGAGCCTGGGCCTGACCCCCAACT
TCAAGAGCAACTTCGACCTGGCCGAGGATGCCAAACTGCAGCTGAGCAAGGACACCTACGACGACGA
CCTGGACAACCTGCTGGCCCAGATCGGCGACCAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGT
CCGACGCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCCTGAGCGCC
TCTATGATCAAGAGATACGACGAGCACCACCAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCA
GCTGCCTGAGAAGTACAAAGAGATTTTCTTCGACCAGAGCAAGAACGGCTACGCCGGCTACATTGACG
GCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATCCTGGAAAAGATGGACGGCACCGA
GGAACTGCTCGTGAAGCTGAAGAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGACAACGGCAGC
ATCCCCCACCAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATT
CCTGAAGGACAACCGGGAAAAGATCGAGAAGATCCTGACCTTCCGCATCCCCTACTACGTGGGCCCTC
TGGCCAGGGGAAACAGCAGATTCGCCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCCTGGAA
CTTCGAGGAAGTGGTGGACAAGGGCGCTTCCGCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATA
AGAACCTGCCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTGTACGAGTACTTCACCGTGTATAAC
GAGCTGACCAAAGTGAAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGCGGCGAGCAGA
AAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGA
CTACTTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCT
CCCTGGGCACATACCACGATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAA
CGAGGACATTCTGGAAGATATCGTGCTGACCCTGACACTGTTTGAGGACAGAGAGATGATCGAGGAA
CGGCTGAAAACCTATGCCCACCTGTTCGACGACAAAGTGATGAAGCAGCTGAAGCGGCGGAGATACA
CCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGCAAGACAAT
CCTGGATTTCCTGAAGTCCGACGGCTTCGCCAACAGAAACTTCATGCAGCTGATCCACGACGACAGCC
TGACCTTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACAT
TGCCAATCTGGCCGGCAGCCCCGCCATTAAGAAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAG
CTCGTGAAAGTGATGGGCCGGCACAAGCCCGAGAACATCGTGATCGAAATGGCCAGAGAGAACCAGA
CCACCCAGAAGGGACAGAAGAACAGCCGCGAGAGAATGAAGCGGATCGAAGAGGGCATCAAAGAGC
TGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTGTACCT
GTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAACTGGACATCAACCGGCTGTCCGACT
ACGATGTGGACGCTATCGTGCCTCAGAGCTTTCTGAAGGACGACTCCATCGACAACAAGGTGCTGACC
AGAAGCGACAAGAACCGGGGCAAGAGCGACAACGTGCCCTCCGAAGAGGTCGTGAAGAAGATGAAG
AACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATTACCCAGAGAAAGTTCGACAATCTGACCAAGG
CCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGACAGCTGGTGGAAACCCG
GCAGATCACAAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAATGAC
AAGCTGATCCGGGAAGTGAAAGTGATCACCCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTT
CCAGTTTTACAAAGTGCGCGAGATCAACAACTACCACCACGCCCACGACGCCTACCTGAACGCCGTCG
TGGGAACCGCCCTGATCAAAAAGTACCCTAAGCTGGAAAGCGAGTTCGTGTACGGCGACTACAAGGT
GTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAGGAAATCGGCAAGGCTACCGCCAAGTACTTC
TTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAACGGCGAGATCCGGAAGCG
GCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCCACC
GTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCT
TCAGCAAAGAGTCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGA
CCCTAAGAAGTACGGCGGCTTCGACAGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGG
AAAAGGGCAAGTCCAAGAAACTGAAGAGTGTGAAAGAGCTGCTGGGGATCACCATCATGGAAAGAA
GCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAAGTGAAAAAGGACCT
GATCATCAAGCTGCCTAAGTACTCCCTGTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCTG
CCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCC
AGCCACTATGAGAAGCTGAAGGGCTCCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGC
ACAAGCACTACCTGGACGAGATCATCGAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGA
CGCTAATCTGGACAAAGTGCTGTCCGCCTACAACAAGCACCGGGATAAGCCCATCAGAGAGCAGGCC
GAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTCAAGTACTTTGACACC
ACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGCA
TCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGTGACTCCGGCGGAAGCTCTGGT
GGCAGCAAGCGGACCGCCGACGGCTCTGAATTCGAGAGCCCTAAGAAGAAAAGAAAGGTGAGCGGA
GGCTCTAGCGGCGGAAGCACCCTGAACATTGAAGACGAGTATAGACTGCATGAAACAAGCAAGGAAC
CCGACGTGTCCCTGGGCTCCACCTGGCTGTCCGACTTTCCCCAGGCCTGGGCCGAGACAGGAGGAATG
GGCCTGGCCGTGCGGCAGGCACCCCTGATCATCCCTCTGAAGGCCACCTCTACACCCGTGAGCATCAA

GCAGTACCCTATGTCTCAGGAGGCCAGACTGGGCATCAAGCCTCACATCCAGAGGCTGCTGGACCAGG
GCATCCTGGTGCCATGCCAGAGCCCCTGGAACACACCACTGCTGCCCGTGAAGAAGCCAGGCACCAAT
GACTATAGACCCGTGCAGGATCTGAGAGAGGTGAACAAGAGGGTGGAGGATATCCACCCCACCGTGC
CCAACCCTTACAATCTGCTGTCCGGCCTGCCCCCTTCTCACCAGTGGTATACAGTGCTGGACCTGAAGG
ATGCCTTCTTTTGTCTGAGACTGCACCCTACCAGCCAGCCACTGTTCGCCTTTGAGTGGAGGGACCCTG
AGATGGGCATCTCTGGCCAGCTGACCTGGACACGCCTGCCTCAGGGCTTCAAGAATAGCCCAACACTG
TTTAACGAGGCCCTGCACCGCGACCTGGCAGATTTCCGGATCCAGCACCCAGATCTGATCCTGCTGCA
GTACGTGGACGATCTGCTGCTGGCCGCCACCAGCGAGCTGGATTGCCAGCAGGGAACACGCGCCCTGC
TGCAGACCCTGGGAAACCTGGGATATAGGGCATCCGCCAAGAAGGCCCAGATCTGTCAGAAGCAGGT
GAAGTACCTGGGCTATCTGCTGAAGGAGGGCCAGAGATGGCTGACAGAGGCCAGGAAGGAGACAGTG
ATGGGCCAGCCAACACCCAAGACCCCAAGACAGCTGAGGGAGTTCCTGGGCAAAGCAGGATTTTGCA
GGCTGTTCATCCCAGGATTCGCAGAGATGGCAGCACCTCTGTACCCACTGACCAAGCCGGGCACCCTG
TTTAATTGGGGCCCTGACCAGCAGAAGGCCTATCAGGAGATCAAGCAGGCCCTGCTGACAGCACCAGC
CCTGGGCCTGCCAGACCTGACCAAGCCTTTCGAGCTGTTTGTGGATGAGAAGCAGGGCTACGCCAAGG
GCGTGCTGACCCAGAAGCTGGGACCATGGAGACGGCCCGTGGCCTATCTGTCCAAGAAGCTGGACCC
AGTGGCAGCAGGATGGCCACCATGCCTGAGGATGGTGGCAGCAATCGCCGTGCTGACAAAGGATGCC
GGCAAGCTGACCATGGGACAGCCACTGGTCATCCTGGCACCACACGCAGTGGAGGCCCTGGTGAAGC
AGCCTCCAGATCGCTGGCTGTCTAACGCCCGGATGACACACTACCAGGCCCTGCTGCTGGACACCGAT
CGCGTGCAGTTTGGCCCTGTGGTGGCCCTGAATCCAGCCACCCTGCTGCCTCTGCCAGAGGAGGGCCT
GCAGCACAACTGTCTGGACATCCTGGCAGAGGCACACGGAACAAGGCCAGACCTGACCGATCAGCCC
CTGCCTGACGCCGATCACACATGGTATACCGATGGAAGCTCCCTGCTGCAGGAGGGCCAGAGGAAGG
CAGGAGCAGCAGTGACCACAGAGACAGAAGTGATCTGGGCCAAGGCCCTGCCAGCAGGCACATCCGC
CCAGCGGGCCGAGCTGATCGCCCTGACCCAGGCCCTGAAGATGGCCGAGGGCAAGAAGCTGAACGTG
TACACAGACTCCAGATATGCCTTCGCCACCGCACACATCCACGGAGAGATCTACAGGCGCCGGGGCTG
GCTGACCTCTGAGGGCAAGGAGATCAAGAACAAGGATGAGATCCTGGCCCTGCTGAAGGCCCTGTTTC
TGCCCAAGCGGCTGAGCATCATCCACTGTCCTGGACACCAGAAGGGACACTCCGCCGAGGCAAGGGG
CAATCGGATGGCCGACCAGGCCGCCAGAAAGGCTGCTATTACTGAAACTCCCGACACTTCCACTCTGC
TGATTGAAAACTCCTCCCCTTCTGGCGGCTCAAAAAGAACCGCCGACGGCAGCGAATTCGAGTCTCCC
AAGAAGAAGAGGAAAGTCGGCTCTGGCCCTGCCGCTAAGAGAGTGAAGCTGGAC

[512] PEmax component sequences of SEQ ID NO: 99:

Bipartite SV40 NLS:

MKRTADGSEFESPKKKRKV (SEQ ID NO: 101)

SpCas9 R221K N394K H840A:

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRY
TRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
STDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLS
KSRKLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYA
GYIDGGASQEEFYKFIKPILEKMDGTEELLVKLKREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFL
KDNREKIEKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNE
KVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECF
DSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVM
KQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDS
LHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKE
LGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSD
KNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHV
AQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYP
KLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVW
DKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVL

VVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLD
KVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDL
SQLGGD (SEQ ID NO: 104)

Linker = (SGGSx2–bipartite SV40 NLS–SGGSx2):

SGGSSGGSKRTADGSEFESPKKKRKVSGGSSGGS (SEQ ID NO: 105)

Genscript codon optimized MMLV RT pentamutant (D200N T306K W313F T330P L603W):

TLNIEDEYRLHETSKEPDVSLGSTWLSDFPQAWAETGGMGLAVRQAPLIIPLKATSTPVSIKQYPMSQEARL
GIKPHIQRLLDQGILVPCQSPWNTPLLPVKKPGTNDYRPVQDLREVNKRVEDIHPTVPNPYNLLSGLPPSHQ
WYTVLDLKDAFFCLRLHPTSQPLFAFEWRDPEMGISGQLTWTRLPQGFKNSPTLFNEALHRDLADFRIQHP
DLILLQYVDDLLLAATSELDCQQGTRALLQTLGNLGYRASAKKAQICQKQVKYLGYLLKEGQRWLTEAR
KETVMGQPTPKTPRQLREFLGKAGFCRLFIPGFAEMAAPLYPLTKPGTLFNWGPDQQKAYQEIKQALLTAP
ALGLPDLTKPFELFVDEKQGYAKGVLTQKLGPWRRPVAYLSKKLDPVAAGWPPCLRMVAAIAVLTKDAG
KLTMGQPLVILAPHAVEALVKQPPDRWLSNARMTHYQALLLDTDRVQFGPVVALNPATLLPLPEEGLQHN
CLDILAEAHGTRPDLTDQPLPDADHTWYTDGSSLLQEGQRKAGAAVTTETEVIWAKALPAGTSAQRAELI
ALTQALKMAEGKKLNVYTDSRYAFATAHIHGEIYRRRGWLTSEGKEIKNKDEILALLKALFLPKRLSIIHCP
GHQKGHSAEARGNRMADQAARKAAITETPDTSTLLIENSSP (SEQ ID NO: 98)

Other linker sequences:

SGGS (SEQ ID NO: 122)

Bipartite SV40 NLS:

KRTADGSEFESPKKKRKV (SEQ ID NO: 140)

Other linker sequences:

GSG (SEQ ID NO: 122)

c-Myc NLS:

PAAKRVKLD (SEQ ID NO: 135)

[513] The modified fusion proteins may comprise any suitable structural configuration. For example, the fusion protein may comprise from the N-terminal to the C-terminal direction, a napDNAbp fused to a polymerase (e.g., a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase, such as a reverse transcriptase). In other embodiments, the fusion protein may comprise from the N-terminal to the C-terminal direction, a polymerase (e.g., a reverse transcriptase) fused to a napDNAbp. The fused domain may optionally be joined by a linker, e.g., an amino acid sequence. In other embodiments, the fusion proteins may comprise the structure NH_2-[napDNAbp]-[polymerase]-COOH; or NH_2-[polymerase]-[napDNAbp]-COOH, wherein each instance of "]-[" indicates the presence of an optional linker sequence. In

embodiments wherein the polymerase is a reverse transcriptase, the fusion proteins may comprise the structure NH_2-[napDNAbp]-[RT]-COOH; or NH_2-[RT]-[napDNAbp]-COOH, wherein each instance of "]-[" indicates the presence of an optional linker sequence.

[514] In various embodiments, the prime editor fusion proteins utilized in the methods and compositions contemplated herein may also include any variants of the above-disclosed sequences having an amino acid sequence that is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to PEmax.

[515] In certain embodiments, linkers may be used to link any of the peptides or peptide domains or moieties of the invention (e.g., a napDNAbp linked or fused to a reverse transcriptase).

[516] The present disclosure contemplates the modification of any Cas9 protein known in the art with one or more of the PEmax mutations described herein (*i.e.*, R221K, N394K, and/or H840A) and the combination of any modified Cas9 protein with one or more of the PEmax architecture features described herein (*e.g.*, the optimized MMLV RT pentamutant, NLS's, linkers, *etc.*).

[517] In some embodiments, the PEmax proteins described herein include any of the following other Cas9 sequences disclosed herein, or variants thereof, which may be further modified with one or more of the mutations described herein at corresponding amino acid positions. The napDNAbp used in the PEmax constructs described herein may include any suitable homologs and/or orthologs or naturally occurring enzymes, such as Cas9. Cas9 homologs and/or orthologs have been described in various species, including, but not limited to, *S. pyogenes* and *S. thermophilus*. The Cas moiety may be configured (e.g., mutagenized, recombinantly engineered, or otherwise obtained from nature) as a nickase, i.e., capable of cleaving only a single strand of the target double-stranded DNA. Additional suitable Cas9 nucleases and sequences will be apparent to those of skill in the art based on this disclosure, and such Cas9 nucleases and sequences include Cas9 sequences from the organisms and loci disclosed in Chylinski, Rhun, and Charpentier, "The tracrRNA and Cas9 families of type II CRISPR-Cas immunity systems" (2013) *RNA Biology* 10:5, 726-737; the entire contents of which are incorporated herein by reference. In some embodiments, a Cas9 nuclease has an inactive (e.g., an inactivated) DNA

cleavage domain; that is, the Cas9 is a nickase. In some embodiments, the Cas9 protein comprises an amino acid sequence that is at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to the amino acid sequence of a Cas9 protein as provided by any one of the Cas9 orthologs provided herein.

Prime editor-MMR inhibitor fusion proteins

[518] The present disclosure contemplates that in some embodiments, the MMR inhibitor (e.g., an antibody or polypeptide inhibitor of an MMR protein, such as an MLH1 dominant negative variant that inhibits MMR) may be linked to a prime editor fusion protein or at least one of the components thereof.

[519] In certain embodiments, the inhibitor domains described herein, e.g., an anti-MLH1 antibody, or an MLH1 dominant negative variant, may also be provided *in cis* by fusing the domain to the prime editor domain. Any of the following structures are contemplated, wherein "]-[" denotes an optional linker:

[520] [napDNAbp]-[reverse transcriptase]-[MLH1 inhibitor];

[521] [reverse transcriptase]-[napDNAbp]-[MLH1 inhibitor];

[522] [MLH1 inhibitor]-[reverse transcriptase]-[napDNAbp];

[523] [MLH1 inhibitor]-[napDNAbp]-[reverse transcriptase],

[524] [napDNAbp]-[MLH1 inhibitor]-[reverse transcriptase]; or

[525] [reverse transcriptase]-[MLH1 inhibitor]-[napDNAbp].

[526] In certain embodiments, the inhibitor domain is an MLH1 dominant negative variant. Any of the following structures are contemplated, wherein "]-[" denotes an optional linker:

[527] [napDNAbp]-[reverse transcriptase]-[MLH1 dominant negative variant];

[528] [reverse transcriptase]-[napDNAbp]-[MLH1 dominant negative variant];

[529] [MLH1 dominant negative variant]-[reverse transcriptase]-[napDNAbp];

[530] [MLH1 dominant negative variant]-[napDNAbp]-[reverse transcriptase];

[531] [napDNAbp]-[MLH1 dominant negative variant]-[reverse transcriptase]; or

[532] [reverse transcriptase]-[MLH1 dominant negative variant]-[napDNAbp].

[533] In certain other embodiments, the inhibitor domains described herein, e.g., an anti-MMR protein antibody, or an MMR protein dominant negative variant, may also be provided *in cis* by fusing the domain to the prime editor domain. Any of the following structures are contemplated, wherein "]-[" denotes an optional linker:

[534] [napDNAbp]-[reverse transcriptase]-[anti-MMR protein inhibitor];

[535] [reverse transcriptase]-[napDNAbp]-[anti-MMR protein inhibitor];

[536] [anti-MMR protein inhibitor]-[reverse transcriptase]-[napDNAbp];

[537] [anti-MMR protein inhibitor]-[napDNAbp]-[reverse transcriptase];

[538] [napDNAbp]-[anti-MMR protein inhibitor]-[reverse transcriptase]; or

[539] [reverse transcriptase]-[anti-MMR protein inhibitor]-[napDNAbp], wherein the MMR protein is any one of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

[540] In certain embodiments, the inhibitor domain is a dominant negative variant of any MMR protein, such as MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA. Any of the following structures are contemplated, wherein "]-[" denotes an optional linker:

[541] [napDNAbp]-[reverse transcriptase]- [dominant negative variant of any MMR protein];

[542] [reverse transcriptase]-[napDNAbp]-[dominant negative variant of any MMR protein];

[543] [dominant negative variant of any MMR protein]-[reverse transcriptase]-[napDNAbp];

[544] [dominant negative variant of any MMR protein napDNAbp]-[reverse transcriptase];

[545] [napDNAbp]-[dominant negative variant of any MMR protein]-[reverse transcriptase]; or [reverse transcriptase]-[dominant negative variant of any MMR protein]-[napDNAbp].

[546] In addition, the MMR inhibitor may be fused to only one of the domains of a prime editor and administered separate from the other prime editor domain. For example, the MMR inhibitor may be fused to the napDNAbp domain, whereby the polymerase domain is provided separately *in trans*. In another, the MMR inhibitor may be fused to the polymerase domain, whereby the napDNAbp domain is provided separately *in trans*.

A. Linkers

[547] As defined above, the term "linker," as used herein, refers to a chemical group or a molecule linking two molecules or moieties, e.g., a binding domain and a cleavage domain of a nuclease. In some embodiments, a linker joins a gRNA binding domain of an RNA-programmable nuclease and the catalytic domain of a polymerase (e.g., a reverse transcriptase). In some embodiments, a linker joins a dCas9 and reverse transcriptase. Typically, the linker is

positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond, thus connecting the two. In some embodiments, the linker is an amino acid or a plurality of amino acids (e.g., a peptide or protein). In some embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 5-100 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated.

[548] The linker may be as simple as a covalent bond, or it may be a polymeric linker many atoms in length. In certain embodiments, the linker is a polypeptide or based on amino acids. In other embodiments, the linker is not peptide-like. In certain embodiments, the linker is a covalent bond (e.g., a carbon-carbon bond, disulfide bond, carbon-heteroatom bond, etc.). In certain embodiments, the linker is a carbon-nitrogen bond of an amide linkage. In certain embodiments, the linker is a cyclic or acyclic, substituted or unsubstituted, branched or unbranched aliphatic or heteroaliphatic linker. In certain embodiments, the linker is polymeric (e.g., polyethylene, polyethylene glycol, polyamide, polyester, etc.). In certain embodiments, the linker comprises a monomer, dimer, or polymer of aminoalkanoic acid. In certain embodiments, the linker comprises an aminoalkanoic acid (e.g., glycine, ethanoic acid, alanine, beta-alanine, 3-aminopropanoic acid, 4-aminobutanoic acid, 5-pentanoic acid, etc.). In certain embodiments, the linker comprises a monomer, dimer, or polymer of aminohexanoic acid (Ahx). In certain embodiments, the linker is based on a carbocyclic moiety (e.g., cyclopentane, cyclohexane). In other embodiments, the linker comprises a polyethylene glycol moiety (PEG). In other embodiments, the linker comprises amino acids. In certain embodiments, the linker comprises a peptide. In certain embodiments, the linker comprises an aryl or heteroaryl moiety. In certain embodiments, the linker is based on a phenyl ring. The linker may include functionalized moieties to facilitate attachment of a nucleophile (e.g., thiol, amino) from the peptide to the linker. Any electrophile may be used as part of the linker. Exemplary electrophiles include, but are not limited to, activated esters, activated amides, Michael acceptors, alkyl halides, aryl halides, acyl halides, and isothiocyanates.

[549] In some other embodiments, the linker comprises the amino acid sequence $(GGGGS)_n$ (SEQ ID NO: 118), $(G)_n$ (SEQ ID NO: 119), $(EAAAK)_n$ (SEQ ID NO: 120), $(GGS)_n$ (SEQ ID

NO: 121), $(SGGS)_n$ (SEQ ID NO: 122), $(XP)_n$ (SEQ ID NO: 123), or any combination thereof, wherein n is independently an integer between 1 and 30, and wherein X is any amino acid. In some embodiments, the linker comprises the amino acid sequence (GGS)n (SEQ ID NO: 121), wherein n is 1, 3, or 7. In some embodiments, the linker comprises the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 124), also referred to as XTEN. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 125). In some embodiments, the linker comprises the amino acid sequence SGGSGGSGGS (SEQ ID NO: 126). In some embodiments, the linker comprises the amino acid sequence SGGS (SEQ ID NO: 127). In other embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGGSSGGS (SEQ ID NO: 128, 60AA).

[550] In certain embodiments, linkers may be used to link any of the peptides or peptide domains or moieties of the invention (e.g., a napDNAbp linked or fused to a reverse transcriptase).

[551] As defined above, the term "linker," as used herein, refers to a chemical group or a molecule linking two molecules or moieties, e.g., a binding domain and a cleavage domain of a nuclease. In some embodiments, a linker joins a gRNA binding domain of an RNA-programmable nuclease and the catalytic domain of a recombinase. In some embodiments, a linker joins a dCas9 and reverse transcriptase. Typically, the linker is positioned between, or flanked by, two groups, molecules, or other moieties and connected to each one via a covalent bond, thus connecting the two. In some embodiments, the linker is an amino acid or a plurality of amino acids (e.g., a peptide or protein). In some embodiments, the linker is an organic molecule, group, polymer, or chemical moiety. In some embodiments, the linker is 5-100 amino acids in length, for example, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 30-35, 35-40, 40-45, 45-50, 50-60, 60-70, 70-80, 80-90, 90-100, 100-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated.

[552] The linker may be as simple as a covalent bond, or it may be a polymeric linker many atoms in length. In certain embodiments, the linker is a polypeptide or based on amino acids. In other embodiments, the linker is not peptide-like. In certain embodiments, the linker is a covalent bond (e.g., a carbon-carbon bond, disulfide bond, carbon-heteroatom bond, etc.). In certain

embodiments, the linker is a carbon-nitrogen bond of an amide linkage. In certain embodiments, the linker is a cyclic or acyclic, substituted or unsubstituted, branched or unbranched aliphatic or heteroaliphatic linker. In certain embodiments, the linker is polymeric (e.g., polyethylene, polyethylene glycol, polyamide, polyester, etc.). In certain embodiments, the linker comprises a monomer, dimer, or polymer of aminoalkanoic acid. In certain embodiments, the linker comprises an aminoalkanoic acid (e.g., glycine, ethanoic acid, alanine, beta-alanine, 3-aminopropanoic acid, 4-aminobutanoic acid, 5-pentanoic acid, etc.). In certain embodiments, the linker comprises a monomer, dimer, or polymer of amino*HEXA*noic acid (Ahx). In certain embodiments, the linker is based on a carbocyclic moiety (e.g., cyclopentane, cyclo*HEXA*ne). In other embodiments, the linker comprises a polyethylene glycol moiety (PEG). In other embodiments, the linker comprises amino acids. In certain embodiments, the linker comprises a peptide. In certain embodiments, the linker comprises an aryl or heteroaryl moiety. In certain embodiments, the linker is based on a phenyl ring. The linker may include functionalized moieties to facilitate attachment of a nucleophile (e.g., thiol, amino) from the peptide to the linker. Any electrophile may be used as part of the linker. Exemplary electrophiles include, but are not limited to, activated esters, activated amides, Michael acceptors, alkyl halides, aryl halides, acyl halides, and isothiocyanates.

[553] In some other embodiments, the linker comprises the amino acid sequence (GGGGS)n (SEQ ID NO: 118), (G)n (SEQ ID NO: 119), $(EAAAK)_n$ (SEQ ID NO: 120), $(GGS)_n$ (SEQ ID NO: 121), (SGGS)n (SEQ ID NO: 122), (XP)n (SEQ ID NO: 123), or any combination thereof, wherein n is independently an integer between 1 and 30, and wherein X is any amino acid. In some embodiments, the linker comprises the amino acid sequence (GGS)n (SEQ ID NO: 121), wherein n is 1, 3, or 7. In some embodiments, the linker comprises the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 124), also referred to as XTEN. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 125). In some embodiments, the linker comprises the amino acid sequence SGGSGGSGGS (SEQ ID NO: 126). In some embodiments, the linker comprises the amino acid sequence SGGS (SEQ ID NO: 127).

[554] In particular, the following linkers can be used in various embodiments to join prime editor domains with one another: GGS (SEQ ID NO: 129); GGSGGS (SEQ ID NO: 130); GGSGGSGGS (SEQ ID NO: 131); SGGSSGGSSGSETPGTSESATPESSGGSSGGSS (SEQ ID

NO: 102); SGSETPGTSESATPES (SEQ ID NO: 124), also referred to as XTEN; SGGSSGGSSGSETPGTSESATPESAGSYPYDVPDYAGSAAPAAKKKKLDGSGSGGSSGGS (SEQ ID NO: 128).

[555] The PE fusion proteins may also comprise various other domains besides the napDNAbp (e.g., Cas9 domain) and the polymerase domain (e.g., RT domain). For example, in the case where the napDNAbp is a Cas9 and the polymerase is a RT, the PE fusion proteins may comprise one or more linkers that join the Cas9 domain with the RT domain. The linkers may also join other functional domains, such as nuclear localization sequences (NLS) or a FEN1 (or other flap endonuclease) to the PE fusion proteins or a domain thereof.

[556] In some embodiments, the PE fusion proteins may comprise an inhibitor of the DNA mismatch repair pathway (*e.g.*, MLH1dn as described herein). In certain embodiments, a PE fusion protein and an inhibitor of the DNA mismatch repair pathway are fused via a linker. In some embodiments, the linker is a self-hydrolyzing linker. Suitable self-hydrolyzing linkers include, but are not limited to, amino acid sequences comprising 2A self-cleaving peptides. 2A self-cleaving peptides are capable of inducing ribosomal skipping during protein translation, resulting in the failure of the ribosome to make a peptide bond between two genes, or two gene fragments. Exemplary 2A self-cleaving peptides that may be used as linkers in the fusion proteins described herein include the amino acid sequences:

T2A - EGRGSLLTCGDVEENPGP (SEQ ID NO: 233)

P2A - ATNFSLLKQAGDVEENPGP (SEQ ID NO: 234)

E2A - QCTNYALLKLAGDVESNPGP (SEQ ID NO: 235)

F2A - VKQTLNFDLLKLAGDVESNPGP (SEQ ID NO: 236)

[557] In certain embodiments, the PE fusion proteins described herein are fused to MLH1dn by a linker comprising the amino acid sequence of SEQ ID NO: 234.

B. Nuclear localization sequence (NLS)

[558] In various embodiments, the PE fusion proteins may comprise one or more nuclear localization sequences (NLS), which help promote translocation of a protein into the cell nucleus. Such sequences are well-known in the art and can include the examples represented by SEQ ID NOs: 1, 101, 103, 133-139.

[559] The NLS examples above are non-limiting. The PE fusion proteins may comprise any known NLS sequence, including any of those described in Cokol *et al.*, "Finding nuclear

localization signals," *EMBO Rep.*, 2000, 1(5): 411-415 and Freitas *et al.*, "Mechanisms and Signals for the Nuclear Import of Proteins," *Current Genomics*, 2009, 10(8): 550-7, each of which are incorporated herein by reference.

[560] In various embodiments, the prime editors and constructs encoding the prime editors utilized in the methods and compositions disclosed herein further comprise one or more, preferably, at least two nuclear localization signals. In certain embodiments, the prime editors comprise at least two NLSs. In embodiments with at least two NLSs, the NLSs can be the same NLSs or they can be different NLSs. In addition, the NLSs may be expressed as part of a fusion protein with the remaining portions of the prime editors. In some embodiments, one or more of the NLSs are bipartite NLSs ("bpNLS"). In certain embodiments, the disclosed fusion proteins comprise two bipartite NLSs. In some embodiments, the disclosed fusion proteins comprise more than two bipartite NLSs.

[561] The location of the NLS fusion can be at the N-terminus, the C-terminus, or within a sequence of a prime editor (e.g., inserted between the encoded napDNAbp component (e.g., Cas9) and a polymerase domain (e.g., a reverse transcriptase domain).

[562] The NLSs may be any known NLS sequence in the art. The NLSs may also be any future-discovered NLSs for nuclear localization. The NLSs also may be any naturally-occurring NLS, or any non-naturally occurring NLS (e.g., an NLS with one or more desired mutations). The term "nuclear localization sequence" or "NLS" refers to an amino acid sequence that promotes import of a protein into the cell nucleus, for example, by nuclear transport. Nuclear localization sequences are known in the art and would be apparent to the skilled artisan. For example, NLS sequences are described in Plank *et al.*, International PCT application PCT/EP2000/011690, filed November 23, 2000, published as WO/2001/038547 on May 31, 2001, the contents of which are incorporated herein by reference. In some embodiments, an NLS comprises the amino acid sequence PKKKRKV (SEQ ID NO: 132), MDSLLMNRRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 1), KRTADGSEFESPKKKRKV (SEQ ID NO: 140), or KRTADGSEFEPKKKRKV (SEQ ID NO: 141). In other embodiments, NLS comprises the amino acid sequences NLSKRPAAIKKAGQAKKKK (SEQ ID NO: 142), PAAKRVKLD (SEQ ID NO: 135), RQRRNELKRSF (SEQ ID NO: 143), NQSSNFGPMKGGNFGGRSSGPYGGGGQYFAKPRNQGGY (SEQ ID NO: 144).

[563] In one aspect of the disclosure, a prime editor may be modified with one or more nuclear localization signals (NLS), preferably at least two NLSs. In certain embodiments, the prime editors are modified with two or more NLSs. The disclosure contemplates the use of any nuclear localization signal known in the art at the time of the disclosure, or any nuclear localization signal that is identified or otherwise made available in the state of the art after the time of the instant filing. A representative nuclear localization signal is a peptide sequence that directs the protein to the nucleus of the cell in which the sequence is expressed. A nuclear localization signal is predominantly basic, can be positioned almost anywhere in a protein's amino acid sequence, generally comprises a short sequence of four amino acids (Autieri & Agrawal, (1998) *J. Biol. Chem.* 273: 14731-37, incorporated herein by reference) to eight amino acids, and is typically rich in lysine and arginine residues (Magin *et al.*, (2000) *Virology* 274: 11-16, incorporated herein by reference). Nuclear localization signals often comprise proline residues. A variety of nuclear localization signals have been identified and have been used to effect transport of biological molecules from the cytoplasm to the nucleus of a cell. See, e.g., Tinland *et al.*, (1992) *Proc. Natl. Acad. Sci. U.S.A.* 89:7442-46; Moede *et al.*, (1999) *FEBS Lett.* 461:229-34, which is incorporated by reference. Translocation is currently thought to involve nuclear pore proteins.

[564] Most NLSs can be classified in three general groups: (i) a monopartite NLS exemplified by the SV40 large T antigen NLS (PKKKRKV (SEQ ID NO: 132)); (ii) a bipartite motif consisting of two basic domains separated by a variable number of spacer amino acids and exemplified by the *Xenopus nucleoplasmin* NLS (KRXXXXXXXXXXKKKK (SEQ ID NO: 145)); and (iii) noncanonical sequences such as M9 of the hnRNP Al protein, the influenza virus nucleoprotein NLS, and the yeast Gal4 protein NLS (Dingwall and Laskey 1991).
Nuclear localization signals appear at various points in the amino acid sequences of proteins. NLS's have been identified at the N-terminus, the C-terminus and in the central region of proteins. Thus, the disclosure provides prime editors that may be modified with one or more NLSs at the C-terminus, the N-terminus, as well as at in internal region of the prime editor. The residues of a longer sequence that do not function as component NLS residues should be selected so as not to interfere, for example tonically or sterically, with the nuclear localization signal itself. Therefore, although there are no strict limits on the composition of an NLS-comprising sequence, in practice, such a sequence can be functionally limited in length and composition.

[565] The present disclosure contemplates any suitable means by which to modify a prime editor to include one or more NLSs. In one aspect, the prime editors may be engineered to express a prime editor protein that is translationally fused at its N-terminus or its C-terminus (or both) to one or more NLSs, i.e., to form a prime editor-NLS fusion construct. In other embodiments, the prime editor-encoding nucleotide sequence may be genetically modified to incorporate a reading frame that encodes one or more NLSs in an internal region of the encoded prime editor. In addition, the NLSs may include various amino acid linkers or spacer regions encoded between the prime editor and the N-terminally, C-terminally, or internally-attached NLS amino acid sequence, e.g., and in the central region of proteins. Thus, the present disclosure also provides for nucleotide constructs, vectors, and host cells for expressing fusion proteins that comprise a prime editor and one or more NLSs.

[566] The prime editors utilized in the methods and compositions described herein may also comprise nuclear localization signals which are linked to a prime editor through one or more linkers, e.g., and polymeric, amino acid, nucleic acid, polysaccharide, chemical, or nucleic acid linker element. The linkers within the contemplated scope of the disclosure are not intended to have any limitations and can be any suitable type of molecule (e.g., polymer, amino acid, polysaccharide, nucleic acid, lipid, or any synthetic chemical linker domain) and be joined to the prime editor by any suitable strategy that effectuates forming a bond (e.g., covalent linkage, hydrogen bonding) between the prime editor and the one or more NLSs.

C. Flap endonucleases (e.g., FEN1)

[567] In various embodiments, the PE fusion proteins may comprise one or more flap endonucleases (e.g., FEN1), which refers to an enzyme that catalyzes the removal of 5′ single strand DNA flaps. These are enzymes that process the removal of 5′ flaps formed during cellular processes, including DNA replication. The prime editing utilized in the methods and compositions described herein may utilize endogenously supplied flap endonucleases or those provided *in trans* to remove the 5′ flap of endogenous DNA formed at the target site during prime editing. Flap endonucleases are known in the art and can be found described in Patel *et al.*, "Flap endonucleases pass 5′-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5′-ends," *Nucleic Acids Research*, 2012, 40(10): 4507-4519 and Tsutakawa *et al.*, "Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily," *Cell*, 2011, 145(2): 198-211 (each of

which are incorporated herein by reference). An exemplary flap endonuclease is FEN1, which can be represented by the following amino acid sequence:

Description	Sequence	SEQ ID NO:
FEN1 Wild type (wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQN EEGETTSHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAE KQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEA SCAALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILQE LGLNQEQFVDLCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPVP ENWLHKEAHQLFLEPEVLDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGV KRLSKSRQGSTQGRLDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFKR GK	146

[568] The flap endonucleases may also include any FEN1 variant, mutant, or other flap endonuclease ortholog, homolog, or variant. Non-limiting FEN1 variant examples are as follows:

Description	Sequence	SEQ ID NO:
FEN1 K168R (relative to FEN1 wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQN EEGETTSHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAE KQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEA SCAALV**R**AGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILQE LGLNQEQFVDLCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPVP ENWLHKEAHQLFLEPEVLDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGV KRLSKSRQGSTQGRLDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFKR GK	147
FEN1 S187A (relative to FEN1 wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQN EEGETTSHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAE KQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEA SCAALVKAGKVYAAATEDMDCLTFG**A**PVLMRHLTASEAKKLPIQEFHLSRILQ ELGLNQEQFVDLCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPV PENWLHKEAHQLFLEPEVLDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGV KRLSKSRQGSTQGRLDDFFKVTGSLSSAKRKEPEPKGSTKKKAKTGAAGKFKR GK	148
FEN1 K354R (relative to FEN1 wt)	MGIQGLAKLIADVAPSAIRENDIKSYFGRKVAIDASMSIYQFLIAVRQGGDVLQN EEGETTSHLMGMFYRTIRMMENGIKPVYVFDGKPPQLKSGELAKRSERRAEAE KQLQQAQAAGAEQEVEKFTKRLVKVTKQHNDECKHLLSLMGIPYLDAPSEAEA SCAALVKAGKVYAAATEDMDCLTFGSPVLMRHLTASEAKKLPIQEFHLSRILQE LGLNQEQFVDLCILLGSDYCESIRGIGPKRAVDLIQKHKSIEEIVRRLDPNKYPVP ENWLHKEAHQLFLEPEVLDPESVELKWSEPNEEELIKFMCGEKQFSEERIRSGV KRLSKSRQGSTQGRLDDFFKVTGSLSSA**R**RKEPEPKGSTKKKAKTGAAGKFKR GK	149
GEN1	MGVNDLWQILEPVKQHIPLRNLGGKTIAVDLSLWVCEAQTVKKMMGSVMKPH LRNLFFRISYLTQMDVKLVFVMEGEPPKLKADVISKRNQSRYGSSGKSWSQKTG RSHFKSVLRECLHMLECLGIPWVQAAGEAEAMCAYLNAGGHVDGCLTNDGDT FLYGAQTVYRNFTMNTKDPHVDCYTMSSIKSKLGLDRDALVGLAILLGCDYLP KGVPGVGKEQALKLIQILKGQSLLQRFNRWNETSCNSSPQLLVTKKLAHCSVCS HPGSPKDHERNGCRLCKSDKYCEPHDYEYCCPCEWHRTEHDRQLSEVENNIKK KACCCEGFPFHEVIQEFLLNKDKLVKVIRYQRPDLLLFQRFTLEKMEWPNHYAC EKLLVLLTHYDMIERKLGSRNSNQLQPIRIVKTRIRNGVHCFEIEWEKPEHYAME DKQHGEFALLTIEEESLFEAAYPEIVAVYQKQKLEIKGKKQKRIKPKENNLPEPD EVMSFQSHMTLKPTCEIFHKQNSKLNSGISPDPTLPQESISASLNSLLLPKNTPCL NAQEQFMSSLRPLAIQQIKAVSKSLISESSQPNTSSHNISVIADLHLSTIDWEGTSF	150

	SNSPAIQRNTFSHDLKSEVESELSAIPDGFENIPEQLSCESERYTANIKKVLDEDSD GISPEEHLLSGITDLCLQDLPLKERIFTKLSYPQDNLQPDVNLKTLSILSVKESCIA NSGSDCTSHLSKDLPGIPLQNESRDSKILKGDQLLQEDYKVNTSVPYSVSNTVV KTCNVRPPNTALDHSRKVDMQTTRKILMKKSVCLDRHSSDEQSAPVFGKAKYT TQRMKHSSQKHNSSHFKESGHNKLSSPKIHIKETEQCVRSYETAENEESCFPDST KSSLSSLQCHKKENNSGTCLDSPLPLRQRLKLRFQST	
ERCC5	MGVQGLWKLLECSGRQVSPEALEGKILAVDISIWLNQALKGVRDRHGNSIENP HLLTLFHRLCKLLFFRIRPIFVFDGDAPLLKKQTLVKRRQRKDLASSDSRKTTEK LLKTFLKRQAIKTAFRSKRDEALPSLTQVRRENDLYVLPPLQEEEKHSSEEEDEK EWQERMNQKQALQEEFFHNPQAIDIESEDFSSLPPEVKHEILTDMKEFTKRRRTL FEAMPEESDDFSQYQLKGLLKKNYLNQHIEHVQKEMNQQHSGHIRRQYEDEGG FLKEVESRRVVSEDTSHYILIKGIQAKTVAEVDSESLPSSSKMHGMSFDVKSSPC EKLKTEKEPDATPPSPRTLLAMQAALLGSSSEEELESENRRQARGRNAPAAVDE GSISPRTLSAIKRALDDDEDVKVCAGDDVQTGGPGAEEMRINSSTENSDEGLKV RDGKGIPFTATLASSSVNSAEEHVASTNEGREPTDSVPKEQMSLVHVGTEAFPIS DESMIKDRKDRLPLESAVVRHSDAPGLPNGRELTPASPTCTNSVSKNETHAEVL EQQNELCPYESKFDSSLLSSDDETKCKPNSASEVIGPVSLQETSSIVSVPSEAVDN VENVVSFNAKEHENFLETIQEQQTTESAGQDLISIPKAVEPMEIDSEESESDGSFIE VQSVISDEELQAEFPETSKPPSEQGEEELVGTREGEAPAESESLLRDNSERDDVD GEPQEAEKDAEDSLHEWQDINLEELETLESNLLAQQNSLKAQKQQQERIAATVT GQMFLESQELLRLFGIPYIQAPMEAEAQCAILDLTDQTSGTITDDSDIWLFGARH VYRNFFNKNKFVEYYQYVDFHNQLGLDRNKLINLAYLLGSDYTEGIPTVGCVT AMEILNEFPGHGLEPLLKFSEWWHEAQKNPKIRPNPHDTKVKKKLRTLQLTPGF PNPAVAEAYLKPVVDDSKGSFLWGKPDLDKIREFCQRYFGWNRTKTDESLFPV LKQLDAQQTQLRIDSFFRLAQQEKEDAKRIKSQRLNRAVTCMLRKEKEAAASEI EAVSVAMEKEFELLDKAKRKTQKRGITNTLEESSSLKRKRLSDSKRKNTCGGFL GETCLSESSDGSSSEDAESSSLMNVQRRTAAKEPKTSASDSQNSVKEAPVKNGG ATTSSSSDSDDDGGKEKMVLVTARSVFGKKRRKLRRARGRKRKT	151

[569] In various embodiments, the prime editor fusion proteins utilized in the methods and compositions contemplated herein may include any flap endonuclease variant of the above-disclosed sequences having an amino acid sequence that is at least about 70% identical, at least about 80% identical, at least about 90% identical, at least about 95% identical, at least about 96% identical, at least about 97% identical, at least about 98% identical, at least about 99% identical, at least about 99.5% identical, or at least about 99.9% identical to any of the above sequences. Other endonucleases that may be utilized by the instant methods to facilitate removal of the 5′ end single strand DNA flap include, but are not limited to (1) trex 2, (2) exo1 endonuclease (e.g., Keijzers *et al.*, *Biosci Rep.* 2015, 35(3): e00206)

Trex 2

[570] 3′ three prime repair exonuclease 2 (TREX2) – human [Accession No. NM_080701]

MSEAPRAETFVFLDLEATGLPSVEPEIAELSLFAVHRSSLENPEHDESGALVLPRVLDKLTLCMCPERPFTAK ASEITGLSSEGLARCRKAGFDGAVVRTLQAFLSRQAGPICLVAHNGFDYDFPLLCAELRRLGARLPRDTVC LDTLPALRGLDRAHSHGTRARGRQGYSLGSLFHRYFRAEPSAAHSAEGDVHTLLLIFLHRAAELLAWADE QARGWAHIEPMYLPPDDPSLEA (SEQ ID NO: 152).

[571] 3′ three prime repair exonuclease 2 (TREX2) – mouse [Accession No. NM_011907]

MSEPPRAETFVFLDLEATGLPNMDPEIAEISLFAVHRSSLENPERDDSGSLVLPRVLDKLTLCMCPERPFTAK
ASEITGLSSESLMHCGKAGFNGAVVRTLQGFLSRQEGPICLVAHNGFDYDFPLLCTELQRLGAHLPQDTVC
LDTLPALRGLDRAHSHGTRAQGRKSYSLASLFHRYFQAEPSAAHSAEGDVHTLLLIFLHRAPELLAWADEQ
ARSWAHIEPMYVPPDGPSLEA (SEQ ID NO: 153).

[572] 3′ three prime repair exonuclease 2 (TREX2) – rat [Accession No. NM_001107580]

MSEPLRAETFVFLDLEATGLPNMDPEIAEISLFAVHRSSLENPERDDSGSLVLPRVLDKLTLCMCPERPFTAK
ASEITGLSSEGLMNCRKAAFNDAVVRTLQGFLSRQEGPICLVAHNGFDYDFPLLCTELQRLGAHLPRDTVC
LDTLPALRGLDRVHSHGTRAQGRKSYSLASLFHRYFQAEPSAAHSAEGDVNTLLLIFLHRAPELLAWADEQ
ARSWAHIEPMYVPPDGPSLEA (SEQ ID NO: 154).

Exo1

[573] Human exonuclease 1 (EXO1) has been implicated in many different DNA metabolic processes, including DNA mismatch repair (MMR), micro-mediated end-joining, homologous recombination (HR), and replication. Human EXO1 belongs to a family of eukaryotic nucleases, Rad2/XPG, which also include FEN1 and GEN1. The Rad2/XPG family is conserved in the nuclease domain through species from phage to human. The EXO1 gene product exhibits both 5′ exonuclease and 5′ flap activity. Additionally, EXO1 contains an intrinsic 5′ RNase H activity. Human EXO1 has a high affinity for processing double stranded DNA (dsDNA), nicks, gaps, pseudo Y structures and can resolve Holliday junctions using its inherent flap activity. Human EXO1 is implicated in MMR and contain conserved binding domains interacting directly with MLH1 and MSH2. EXO1 nucleolytic activity is positively stimulated by PCNA, MutSα (MSH2/MSH6 complex), 14-3-3, MRN and 9-1-1 complex.

[574] exonuclease 1 (EXO1) Accession No. NM_003686 (Homo sapiens exonuclease 1 (EXO1), transcript variant 3) – isoform A

MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPTDRYVGFCMKFVNMLLSH
GIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLREGKVSEARECFTRSINITHAMAHKVIKAARSQG
VDCLVAPYEADAQLAYLNKAGIVQAIITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFT
EEKFRYMCILSGCDYLSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANNTFLY
QLVFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQIDDYNPDTAMPAHSRSHSW
DDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPSTVGVERVISTKGLNLPRKSSIVKRPRSAELSED
DLLSQYSLSFTKKTKKNSSEGNKSLSFSEVFVPDLVNGPTNKKSVSTPPRTRNKFATFLQRKNEESGAVVVP
GTRSRFFCSSDSTDCVSNKVSIQPLDETAVTDKENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIPGDH
IPDKATVFTDEESYSFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTALQQFRRKSDSPTSLPENN
MSDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSNASKLSQCSSKDSDSEESDCNIKLLDSQSDQTS
KLRLSHFSKKDTPLRNKVPGLYKSSSADSLSTTKIKPLGPARASGLSKKPASIQKRKHHNAENKPGLQIKLN
ELWKNFGFKKF (SEQ ID NO: 155).

[575] exonuclease 1 (EXO1) Accession No. NM_006027 (Homo sapiens exonuclease 1 (EXO1), transcript variant 3) – isoform B

MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPTDRYVGFCMKFVNMLLSH
GIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLREGKVSEARECFTRSINITHAMAHKVIKAARSQG
VDCLVAPYEADAQLAYLNKAGIVQAIITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFT
EEKFRYMCILSGCDYLSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANNTFLY
QLVFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQIDDYNPDTAMPAHSRSHSW
DDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPSTVGVERVISTKGLNLPRKSSIVKRPRSAELSED
DLLSQYSLSFTKKTKKNSSEGNKSLSFSEVFVPDLVNGPTNKKSVSTPPRTRNKFATFLQRKNEESGAVVVP
GTRSRFFCSSDSTDCVSNKVSIQPLDETAVTDKENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIPGDH
IPDKATVFTDEESYSFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTALQQFRRKSDSPTSLPENN
MSDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSNASKLSQCSSKDSDSEESDCNIKLLDSQSDQTS
KLRLSHFSKKDTPLRNKVPGLYKSSSADSLSTTKIKPLGPARASGLSKKPASIQKRKHHNAENKPGLQIKLN
ELWKNFGFKKDSEKLPPCKKPLSPVRDNIQLTPEAEEDIFNKPECGRVQRAIFQ (SEQ ID NO: 156).

[576] exonuclease 1 (EXO1) Accession No. NM_001319224 (Homo sapiens exonuclease 1 (EXO1), transcript variant 4) – isoform C

MGIQGLLQFIKEASEPIHVRKYKGQVVAVDTYCWLHKGAIACAEKLAKGEPTDRYVGFCMKFVNMLLSH
GIKPILVFDGCTLPSKKEVERSRRERRQANLLKGKQLLREGKVSEARECFTRSINITHAMAHKVIKAARSQG
VDCLVAPYEADAQLAYLNKAGIVQAIITEDSDLLAFGCKKVILKMDQFGNGLEIDQARLGMCRQLGDVFT
EEKFRYMCILSGCDYLSSLRGIGLAKACKVLRLANNPDIVKVIKKIGHYLKMNITVPEDYINGFIRANNTFLY
QLVFDPIKRKLIPLNAYEDDVDPETLSYAGQYVDDSIALQIALGNKDINTFEQIDDYNPDTAMPAHSRSHSW
DDKTCQKSANVSSIWHRNYSPRPESGTVSDAPQLKENPSTVGVERVISTKGLNLPRKSSIVKRPRSELSEDD
LLSQYSLSFTKKTKKNSSEGNKSLSFSEVFVPDLVNGPTNKKSVSTPPRTRNKFATFLQRKNEESGAVVVPG
TRSRFFCSSDSTDCVSNKVSIQPLDETAVTDKENNLHESEYGDQEGKRLVDTDVARNSSDDIPNNHIPGDHI
PDKATVFTDEESYSFESSKFTRTISPPTLGTLRSCFSWSGGLGDFSRTPSPSPSTALQQFRRKSDSPTSLPENN
MSDVSQLKSEESSDDESHPLREEACSSQSQESGEFSLQSSNASKLSQCSSKDSDSEESDCNIKLLDSQSDQTS
KLRLSHFSKKDTPLRNKVPGLYKSSSADSLSTTKIKPLGPARASGLSKKPASIQKRKHHNAENKPGLQIKLN
ELWKNFGFKKDSEKLPPCKKPLSPVRDNIQLTPEAEEDIFNKPECGRVQRAIFQ (SEQ ID NO: 157).

D. Inteins and split-inteins

[577] It will be understood that in some embodiments (e.g., delivery of a prime editor *in vivo* using AAV particles), it may be advantageous to split a polypeptide (e.g., a deaminase or a napDNAbp) or a fusion protein (e.g., a prime editor) into an N-terminal half and a C-terminal half, delivery them separately, and then allow their colocalization to reform the complete protein (or fusion protein as the case may be) within the cell. Separate halves of a protein or a fusion protein may each comprise a split-intein tag to facilitate the reformation of the complete protein or fusion protein by the mechanism of protein trans splicing.

[578] Protein trans-splicing, catalyzed by split inteins, provides an entirely enzymatic method for protein ligation. A split-intein is essentially a contiguous intein (e.g., a mini-intein) split into two pieces named N-intein and C-intein, respectively. The N-intein and C-intein of a split intein can associate non-covalently to form an active intein and catalyze the splicing reaction essentially in same way as a contiguous intein does. Split inteins have been found in nature and also engineered in laboratories. As used herein, the term "split intein" refers to any intein in

which one or more peptide bond breaks exists between the N-terminal and C-terminal amino acid sequences such that the N-terminal and C-terminal sequences become separate molecules that can non-covalently reassociate, or reconstitute, into an intein that is functional for trans-splicing reactions. Any catalytically active intein, or fragment thereof, may be used to derive a split intein for use in the methods of the invention. For example, in one aspect the split intein may be derived from a eukaryotic intein. In another aspect, the split intein may be derived from a bacterial intein. In another aspect, the split intein may be derived from an archaeal intein. Preferably, the split intein so-derived will possess only the amino acid sequences essential for catalyzing trans-splicing reactions.

[579] As used herein, the "N-terminal split intein (In)" refers to any intein sequence that comprises an N- terminal amino acid sequence that is functional for trans-splicing reactions. An In thus also comprises a sequence that is spliced out when trans-splicing occurs. An In can comprise a sequence that is a modification of the N-terminal portion of a naturally occurring intein sequence. For example, an In can comprise additional amino acid residues and/or mutated residues so long as the inclusion of such additional and/or mutated residues does not render the In non-functional in trans-splicing. Preferably, the inclusion of the additional and/or mutated residues improves or enhances the trans-splicing activity of the In.

[580] As used herein, the "C-terminal split intein (Ic)" refers to any intein sequence that comprises a C- terminal amino acid sequence that is functional for trans-splicing reactions. In one aspect, the Ic comprises 4 to 7 contiguous amino acid residues, at least 4 amino acids of which are from the last β-strand of the intein from which it was derived. An Ic thus also comprises a sequence that is spliced out when trans-splicing occurs. An Ic can comprise a sequence that is a modification of the C-terminal portion of a naturally occurring intein sequence. For example, an Ic can comprise additional amino acid residues and/or mutated residues so long as the inclusion of such additional and/or mutated residues does not render the In non-functional in trans-splicing. Preferably, the inclusion of the additional and/or mutated residues improves or enhances the trans-splicing activity of the Ic.

[581] In some embodiments of the invention, a peptide linked to an Ic or an In can comprise an additional chemical moiety including, among others, fluorescence groups, biotin, polyethylene glycol (PEG), amino acid analogs, unnatural amino acids, phosphate groups, glycosyl groups, radioisotope labels, and pharmaceutical molecules. In other embodiments, a peptide linked to an

Ic can comprise one or more chemically reactive groups including, among others, ketone, aldehyde, Cys residues and Lys residues. The N-intein and C-intein of a split intein can associate non-covalently to form an active intein and catalyze the splicing reaction when an "intein-splicing polypeptide (ISP)" is present. As used herein, "intein-splicing polypeptide (ISP)" refers to the portion of the amino acid sequence of a split intein that remains when the Ic, In, or both, are removed from the split intein. In certain embodiments, the In comprises the ISP. In another embodiment, the Ic comprises the ISP. In yet another embodiment, the ISP is a separate peptide that is not covalently linked to In nor to Ic.

[582] Split inteins may be created from contiguous inteins by engineering one or more split sites in the unstructured loop or intervening amino acid sequence between the -12 conserved beta-strands found in the structure of mini-inteins. Some flexibility in the position of the split site within regions between the beta-strands may exist, provided that creation of the split will not disrupt the structure of the intein, the structured beta-strands in particular, to a sufficient degree that protein splicing activity is lost.

[583] In protein trans-splicing, one precursor protein consists of an N-extein part followed by the N-intein, another precursor protein consists of the C-intein followed by a C-extein part, and a trans-splicing reaction (catalyzed by the N- and C-inteins together) excises the two intein sequences and links the two extein sequences with a peptide bond. Protein trans-splicing, being an enzymatic reaction, can work with very low (e.g., micromolar) concentrations of proteins and can be carried out under physiological conditions.

[584] Exemplary sequences are as follows:

NAME	SEQUENCE OF LIGAND-DEPENDENT INTEIN
2-4 INTEIN:	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFDQGTRDVIGLRIA GGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGSGNSLALSLTADQMVSALLDAEP PILYSEYDPTSPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLE ILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQRLAQLL LILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALD DKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHNC (SEQ ID NO: 158)
3-2 INTEIN	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVSWFDQGTRDVIGLRIA GGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGSGNSLALSLTADQMVSALLDAEP PILYSEYDPTSPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLE ILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQRLAQLL LILSHIRHMSNKGMEHLYSMKYTNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALDD KFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHNC (SEQ ID NO: 159)

30R3-1 INTEIN	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFDQGTRDVIGLRIA GGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGSGNSLALSLTADQMVSALLDAEP PIPYSEYDPTSPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLE ILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQRLAQLL LILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALD DKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHNC (SEQ ID NO: 160)
30R3-2 INTEIN	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFDQGTRDVIGLRIA GGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGSGNSLALSLTADQMVSALLDAEP PILYSEYDPTSPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLE ILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQRLAQLL LILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALD DKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHNC (SEQ ID NO: 161)
30R3-3 INTEIN	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFDQGTRDVIGLRIA GGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGSGNSLALSLTADQMVSALLDAEP PIPYSEYDPTSPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLECAWLE ILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQRLAQLL LILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALD DKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHNC (SEQ ID NO: 162)
37R3-1 INTEIN	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFDQGTRDVIGLRIA GGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGSGNSLALSLTADQMVSALLDAEP PILYSEYNPTSPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLE ILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQRLAQLL LILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALD DKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHNC ((SEQ ID NO: 163)
37R3-2 INTEIN	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAAAKDGTLLARPVVSWFDQGTRDVIGLRIA GGAIVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGSGNSLALSLTADQMVSALLDAEP PILYSEYDPTSPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLE ILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQRLAQLL LILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALD DKFLHDMLAEGLRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHNC (SEQ ID NO: 164)
37R3-3 INTEIN	CLAEGTRIFDPVTGTTHRIEDVVDGRKPIHVVAVAKDGTLLARPVVSWFDQGTRDVIGLRIA GGATVWATPDHKVLTEYGWRAAGELRKGDRVAGPGGSGNSLALSLTADQMVSALLDAEP PILYSEYDPTSPFSEASMMGLLTNLADRELVHMINWAKRVPGFVDLTLHDQAHLLERAWLE ILMIGLVWRSMEHPGKLLFAPNLLLDRNQGKCVEGMVEIFDMLLATSSRFRMMNLQGEEF VCLKSIILLNSGVYTFLSSTLKSLEEKDHIHRALDKITDTLIHLMAKAGLTLQQQHQRLAQLL LILSHIRHMSNKGMEHLYSMKYKNVVPLYDLLLEMLDAHRLHAGGSGASRVQAFADALD DKFLHDMLAEELRYSVIREVLPTRRARTFDLEVEELHTLVAEGVVVHNC (SEQ ID NO: 165)

[585] Although inteins are most frequently found as a contiguous domain, some exist in a naturally split form. In this case, the two fragments are expressed as separate polypeptides and must associate before splicing takes place, so-called protein trans-splicing.

[586] An exemplary split intein is the *Ssp* DnaE intein, which comprises two subunits, namely, DnaE-N and DnaE-C. The two different subunits are encoded by separate genes, namely *dnaE-n* and *dnaE-c*, which encode the DnaE-N and DnaE-C subunits, respectively. DnaE is a naturally

occurring split intein in *Synechocytis sp.* PCC6803 and is capable of directing trans-splicing of two separate proteins, each comprising a fusion with either DnaE-N or DnaE-C.

[587] Additional naturally occurring or engineered split-intein sequences are known in the or can be made from whole-intein sequences described herein or those available in the art. Examples of split-intein sequences can be found in Stevens *et al.*, "A promiscuous split intein with expanded protein engineering applications," PNAS, 2017, Vol.114: 8538-8543; Iwai *et al.*, "Highly efficient protein trans-splicing by a naturally split DnaE intein from Nostc punctiforme, FEBS Lett, 580: 1853-1858, each of which are incorporated herein by reference. Additional split intein sequences can be found, for example, in WO 2013/045632, WO 2014/055782, WO 2016/069774, and EP2877490, the contents each of which are incorporated herein by reference. In addition, protein splicing in trans has been described *in vivo* and *in vitro* (Shingledecker, *et al.*, *Gene* 207:187 (1998), Southworth, *et al.*, *EMBO J.* 17:918 (1998); Mills, *et al.*, *Proc. Natl. Acad. Sci. USA,* 95:3543-3548 (1998); Lew, *et al.*, *J. Biol. Chem.,* 273:15887-15890 (1998); Wu, *et al.*, *Biochim. Biophys. Acta* 35732:1 (1998b), Yamazaki, *et al.*, *J. Am. Chem. Soc.* 120:5591 (1998), Evans, *et al.*, *J. Biol. Chem.* 275:9091 (2000); Otomo, *et al.*, *Biochemistry* 38:16040-16044 (1999); Otomo, *et al.*, *J. Biolmol. NMR* 14:105-114 (1999); Scott, *et al.*, *Proc. Natl. Acad. Sci. USA* 96:13638-13643 (1999)) and provides the opportunity to express a protein as to two inactive fragments that subsequently undergo ligation to form a functional product.

RNA-protein interaction domain

[588] In various embodiments, two separate protein domains (e.g., a Cas9 domain and a polymerase domain) may be colocalized to one another to form a functional complex (akin to the function of a fusion protein comprising the two separate protein domains) by using an "RNA-protein recruitment system," such as the "MS2 tagging technique." Such systems generally tag one protein domain with an "RNA-protein interaction domain" (aka "RNA-protein recruitment domain") and the other with an "RNA-binding protein" that specifically recognizes and binds to the RNA-protein interaction domain, e.g., a specific hairpin structure. These types of systems can be leveraged to colocalize the domains of a prime editor, as well as to recruitment additional functionalities to a prime editor, such as a UGI domain. In one example, the MS2 tagging technique is based on the interaction of the MS2 bacteriophage coat protein ("MCP" or "MS2cp") with a stem-loop or hairpin structure present in the genome of the phage, i.e., the "MS2 hairpin." In the case of the MS2 hairpin, it is recognized and bound by the MS2

bacteriophage coat protein (MCP). Thus, in one exemplary scenario a deaminase-MS2 fusion can recruit a Cas9-MCP fusion.

[589] A review of other modular RNA-protein interaction domains are described in the art, for example, in Johansson *et al.*, "RNA recognition by the MS2 phage coat protein," *Sem Virol.*, 1997, Vol. 8(3): 176-185; Delebecque *et al.*, "Organization of intracellular reactions with rationally designed RNA assemblies," *Science*, 2011, Vol. 333: 470-474; Mali *et al.*, "Cas9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering," *Nat. Biotechnol.*, 2013, Vol.31: 833-838; and Zalatan *et al.*, "Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds," *Cell*, 2015, Vol.160: 339-350, each of which are incorporated herein by reference in their entireties. Other systems include the PP7 hairpin, which specifically recruits the PCP protein, and the "com" hairpin, which specifically recruits the Com protein. See Zalatan *et al.*

[590] The nucleotide sequence of the MS2 hairpin (or equivalently referred to as the "MS2 aptamer") is: GCCAACATGAGGATCACCCATGTCTGCAGGGCC (SEQ ID NO: 166). The amino acid sequence of the MCP or MS2cp is: GSASNFTQFVLVDNGGTGDVTVAPSNFANGVAEWISSNSRSQAYKVTCSVRQSSAQNR KYTIKVEVPKVATQTVGGEELPVAGWRSYLNMELTIPIFATNSDCELIVKAMQGLLKDG NPIPSAIAANSGIY (SEQ ID NO: 167).

E. UGI domain

[591] In other embodiments, the prime editors utilized in the methods and compositions described herein may comprise one or more uracil glycosylase inhibitor domains. The term "uracil glycosylase inhibitor (UGI)" or "UGI domain," as used herein, refers to a protein that is capable of inhibiting a uracil-DNA glycosylase base-excision repair enzyme. In some embodiments, a UGI domain comprises a wild-type UGI or a UGI as set forth in SEQ ID NO: 168. In some embodiments, the UGI proteins provided herein include fragments of UGI and proteins homologous to a UGI or a UGI fragment. For example, in some embodiments, a UGI domain comprises a fragment of the amino acid sequence set forth in SEQ ID NO: 168. In some embodiments, a UGI fragment comprises an amino acid sequence that comprises at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% of the amino acid sequence as set forth in SEQ ID NO: 168. In some embodiments, a UGI comprises an amino acid sequence

homologous to the amino acid sequence set forth in SEQ ID NO: 168, or an amino acid sequence homologous to a fragment of the amino acid sequence set forth in SEQ ID NO: 168. In some embodiments, proteins comprising UGI or fragments of UGI or homologs of UGI or UGI fragments are referred to as "UGI variants." A UGI variant shares homology to UGI, or a fragment thereof. For example, a UGI variant is at least 70% identical, at least 75% identical, at least 80% identical, at least 85% identical, at least 90% identical, at least 95% identical, at least 96% identical, at least 97% identical, at least 98% identical, at least 99% identical, at least 99.5% identical, or at least 99.9% identical to a wild type UGI or a UGI as set forth in SEQ ID NO: 168. In some embodiments, the UGI variant comprises a fragment of UGI, such that the fragment is at least 70% identical, at least 80% identical, at least 90% identical, at least 95% identical, at least 96% identical, at least 97% identical, at least 98% identical, at least 99% identical, at least 99.5% identical, or at least 99.9% to the corresponding fragment of wild-type UGI or a UGI as set forth in SEQ ID NO: 168. In some embodiments, the UGI comprises the following amino acid sequence:

Uracil-DNA glycosylase inhibitor:

>sp|P14739|UNGI_BPPB2
MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLTSDAPEYKPWALVIQ
DSNGENKIKML (SEQ ID NO: 168).

[592] The prime editors utilized in the methods and compositions described herein may comprise more than one UGI domain, which may be separated by one or more linkers as described herein.

F. Additional PE elements

[593] In certain embodiments, the prime editors utilized in the methods and compositions described herein may comprise an inhibitor of base repair. The term "inhibitor of base repair" or "IBR" refers to a protein that is capable in inhibiting the activity of a nucleic acid repair enzyme, for example a base excision repair enzyme. In some embodiments, the IBR is an inhibitor of OGG base excision repair. In some embodiments, the IBR is an inhibitor of base excision repair ("iBER"). Exemplary inhibitors of base excision repair include inhibitors of APE1, Endo III, Endo IV, Endo V, Endo VIII, Fpg, hOGG1, hNEIL1, T7 EndoI, T4PDG, UDG, hSMUG1, and hAAG. In some embodiments, the IBR is an inhibitor of Endo V or hAAG. In some embodiments, the IBR is an iBER that may be a catalytically inactive glycosylase or catalytically

inactive dioxygenase or a small molecule or peptide inhibitor of an oxidase, or variants threreof. In some embodiments, the IBR is an iBER that may be a TDG inhibitor, MBD4 inhibitor or an inhibitor of an AlkBH enzyme. In some embodiments, the IBR is an iBER that comprises a catalytically inactive TDG or catalytically inactive MBD4. An exemplary catalytically inactive TDG is an N140A mutant of SEQ ID NO: 172 (human TDG).

[594] Some exemplary glycosylases are provided below. The catalytically inactivated variants of any of these glycosylase domains are iBERs that may be fused to the napDNAbp or polymerase domain of the prime editors utilized in the methods and compositions provided in this disclosure.

[595] OGG (human)

MPARALLPRRMGHRTLASTPALWASIPCPRSELRLDLVLPSGQSFRWREQSPAHWSGVLADQVWTLTQTE EQLHCTVYRGDKSQASRPTPDELEAVRKYFQLDVTLAQLYHHWGSVDSHFQEVAQKFQGVRLLRQDPIEC LFSFICSSNNNIARITGMVERLCQAFGPRLIQLDDVTYHGFPSLQALAGPEVEAHLRKLGLGYRARYVSASA RAILEEQGGLAWLQQLRESSYEEAHKALCILPGVGTKVADCICLMALDKPQAVPVDVHMWHIAQRDYSW HPTTSQAKGPSPQTNKELGNFFRSLWGPYAGWAQAVLFSADLRQSRHAQEPPAKRRKGSKGPEG (SEQ ID NO: 169)

[596] MPG (human)

MVTPALQMKKPKQFCRRMGQKKQRPARAGQPHSSSDAAQAPAEQPHSSSDAAQAPCPRERCLGPPTTPGP YRSIYFSSPKGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELRGRIVETEAYLGPEDEAAHSRGGRQ TPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDGACVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDR ELCSGPSKLCQALAINKSFDQRDLAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGS PWVSVVDRVAEQDTQA (SEQ ID NO: 170)

[597] MBD4 (human)

MGTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRSSECNPLLQEPIAS AQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQGLKFRSKSSLANYLHKNGETSLKPEDFD FTVLSKRGIKSRYKDCSMAALTSHLQNQSNNSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLL LKEDEGVDDVNFRKVRKPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQL DRTVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHNEKYEDTFLESEEI GTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTIPRTQIERRKTSLYFSSKYNKEALSPPR RKAFKKWTPPRSPFNLVQETLFHDPWKLLIATIFLNRTSGKMAIPVLWKFLEKYPSAEVARTADWRDVSEL LKPLGLYDLRAKTIVKFSDEYLTKQWKYPIELHGIGKYGNDSYRIFCVNEWKQVHPEDHKLNKYHDWLW ENHEKLSLS (SEQ ID NO: 171)

TDG (human)

MEAENAGSYSLQQAQAFYTFPFQQLMAEAPNMAVVNEQQMPEEVPAPAPAQEPVQEAPKGRKRKPRTTE PKQPVEPKKPVESKKSGKSAKSKEKQEKITDTFKVKRKVDRFNGVSEAELLTKTLPDILTFNLDIVIIGINPGL MAAYKGHHYPGPGNHFWKCLFMSGLSEVQLNHMDDHTLPGKYGIGFTNMVERTTPGSKDLSSKEFREGG RILVQKLQKYQPRIAVFNGKCIYEIFSKEVFGVKVKNLEFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQD KVHYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQLAQEDAKKMAVKEEKYDPGYEAAYGGAYGENPCS SEPCGFSSNGLIESVELRGESAFSGIPNGQWMTQSFTDQIPSFSNHCGTQEQEEESHA (SEQ ID NO: 172)

[598] In some embodiments, the fusion proteins described herein may comprise one or more heterologous protein domains (e.g., about or more than about 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, or more domains in addition to the prime editor components). A fusion protein may comprise any additional protein sequence, and optionally a linker sequence between any two domains. Other exemplary features that may be present are localization sequences, such as cytoplasmic localization sequences, export sequences, such as nuclear export sequences, or other localization sequences, as well as sequence tags that are useful for solubilization, purification, or detection of the fusion proteins.

[599] Examples of protein domains that may be fused to a prime editor or component thereof (e.g., the napDNAbp domain, the polymerase domain, or the NLS domain) include, without limitation, epitope tags, and reporter gene sequences. Non-limiting examples of epitope tags include histidine (His) tags, V5 tags, FLAG tags, influenza hemagglutinin (HA) tags, Myc tags, VSV-G tags, and thioredoxin (Trx) tags. Examples of reporter genes include, but are not limited to, glutathione-5-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT), beta-galactosidase, beta-glucuronidase, luciferase, green fluorescent protein (GFP), HcRed, DsRed, cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and autofluorescent proteins including blue fluorescent protein (BFP). A prime editor may be fused to a gene sequence encoding a protein or a fragment of a protein that bind DNA molecules or bind other cellular molecules, including, but not limited to, maltose binding protein (MBP), S-tag, Lex A DNA binding domain (DBD) fusions, GAL4 DNA binding domain fusions, and herpes simplex virus (HSV) BP16 protein fusions. Additional domains that may form part of a prime editor are described in US Patent Publication No. 2011/0059502, published March 10, 2011 and incorporated herein by reference in its entirety.

[600] In an aspect of the disclosure, a reporter gene which includes, but is not limited to, glutathione-5-transferase (GST), horseradish peroxidase (HRP), chloramphenicol acetyltransferase (CAT) beta-galactosidase, beta-glucuronidase, luciferase, green fluorescent protein (GFP), HcRed, DsRed, cyan fluorescent protein (CFP), yellow fluorescent protein (YFP), and autofluorescent proteins including blue fluorescent protein (BFP), may be introduced into a cell to encode a gene product which serves as a marker by which to measure the alteration or modification of expression of the gene product. In certain embodiments of the disclosure the

gene product is luciferase. In a further embodiment of the disclosure the expression of the gene product is decreased.

[601] Suitable protein tags provided herein include, but are not limited to, biotin carboxylase carrier protein (BCCP) tags, myc-tags, calmodulin-tags, FLAG-tags, hemagglutinin (HA)-tags, polyhistidine tags, also referred to as histidine tags or His-tags, maltose binding protein (MBP)-tags, nus-tags, glutathione-S-transferase (GST)-tags, green fluorescent protein (GFP)-tags, thioredoxin-tags, S-tags, Softags (e.g., Softag 1, Softag 3), strep-tags , biotin ligase tags, FlAsH tags, V5 tags, and SBP-tags. Additional suitable sequences will be apparent to those of skill in the art. In some embodiments, the fusion protein comprises one or more His tags.

[602] In some embodiments of the present disclosure, the activity of the prime editing system may be temporally regulated by adjusting the residence time, the amount, and/or the activity of the expressed components of the PE system. For example, as described herein, the PE may be fused with a protein domain that is capable of modifying the intracellular half-life of the PE. In certain embodiments involving two or more vectors (e.g., a vector system in which the components described herein are encoded on two or more separate vectors), the activity of the PE system may be temporally regulated by controlling the timing in which the vectors are delivered. For example, in some embodiments a vector encoding the nuclease system may deliver the PE prior to the vector encoding the template. In other embodiments, the vector encoding the PEgRNA may deliver the guide prior to the vector encoding the PE system. In some embodiments, the vectors encoding the PE system and PEgRNA are delivered simultaneously. In certain embodiments, the simultaneously delivered vectors temporally deliver, e.g., the PE, PEgRNA, and/or second strand guide RNA components. In further embodiments, the RNA (such as, e.g., the nuclease transcript) transcribed from the coding sequence on the vectors may further comprise at least one element that is capable of modifying the intracellular half-life of the RNA and/or modulating translational control. In some embodiments, the half-life of the RNA may be increased. In some embodiments, the half-life of the RNA may be decreased. In some embodiments, the element may be capable of increasing the stability of the RNA. In some embodiments, the element may be capable of decreasing the stability of the RNA. In some embodiments, the element may be within the 3' UTR of the RNA. In some embodiments, the element may include a polyadenylation signal (PA). In some embodiments, the element may include a cap, e.g., an upstream mRNA or PEgRNA end. In some embodiments, the RNA may

comprise no PA such that it is subject to quicker degradation in the cell after transcription. In some embodiments, the element may include at least one AU-rich element (ARE). The AREs may be bound by ARE binding proteins (ARE-BPs) in a manner that is dependent upon tissue type, cell type, timing, cellular localization, and environment. In some embodiments the destabilizing element may promote RNA decay, affect RNA stability, or activate translation. In some embodiments, the ARE may comprise 50 to 150 nucleotides in length. In some embodiments, the ARE may comprise at least one copy of the sequence AUUUA. In some embodiments, at least one ARE may be added to the 3' UTR of the RNA. In some embodiments, the element may be a Woodchuck Hepatitis Virus (WHP).

[603] Posttranscriptional Regulatory Element (WPRE), which creates a tertiary structure to enhance expression from the transcript. In further embodiments, the element is a modified and/or truncated WPRE sequence that is capable of enhancing expression from the transcript, as described, for example in Zufferey *et al.*, J Virol, 73(4): 2886-92 (1999) and Flajolet *et al.*, J Virol, 72(7): 6175-80 (1998). In some embodiments, the WPRE or equivalent may be added to the 3' UTR of the RNA. In some embodiments, the element may be selected from other RNA sequence motifs that are enriched in either fast- or slow-decaying transcripts.

In some embodiments, the vector encoding the PE or the PEgRNA may be self-destroyed via cleavage of a target sequence present on the vector by the PE system. The cleavage may prevent continued transcription of a PE or a PEgRNA from the vector. Although transcription may occur on the linearized vector for some amount of time, the expressed transcripts or proteins subject to intracellular degradation will have less time to produce off-target effects without continued supply from expression of the encoding vectors.

PEgRNAs

[604] The prime editing system utilized in the methods and compositions described herein contemplates the use of any suitable PEgRNAs.

PEgRNA architecture

[605] In some embodiments, an extended guide RNA usable in the prime editing system utilized in the methods and compositions disclosed herein whereby a traditional guide RNA includes a ~20 nt protospacer sequence and a gRNA core region, which binds with the napDNAbp. In this embodiment, the guide RNA includes an extended RNA segment at the 5′ end, i.e., a 5′ extension. In this embodiment, the 5′extension includes a reverse transcription

template sequence, a reverse transcription primer binding site, and an optional 5-20 nucleotide linker sequence. The RT primer binding site hybridizes to the free 3′ end that is formed after a nick is formed in the non-target strand of the R-loop, thereby priming reverse transcriptase for DNA polymerization in the 5′-3′ direction.

[606] In another embodiment, an extended guide RNA usable in the prime editing system utilized in the methods and compositions disclosed herein whereby a traditional guide RNA includes a ~20 nt protospacer sequence and a gRNA core, which binds with the napDNAbp. In this embodiment, the guide RNA includes an extended RNA segment at the 3′ end, i.e., a 3′ extension. In this embodiment, the 3′extension includes a reverse transcription template sequence, and a reverse transcription primer binding site. The RT primer binding site hybridizes to the free 3′ end that is formed after a nick is formed in the non-target strand of the R-loop, thereby priming reverse transcriptase for DNA polymerization in the 5′-3′ direction.

[607] In another embodiment, an extend guide RNA usable in the prime editing system utilized in the methods and compositions disclosed herein whereby a traditional guide RNA includes a ~20 nt protospacer sequence and a gRNA core, which binds with the napDNAbp. In this embodiment, the guide RNA includes an extended RNA segment at an intermolecular position within the gRNA core, i.e., an intramolecular extension. In this embodiment, the intramolecular extension includes a reverse transcription template sequence, and a reverse transcription primer binding site. The RT primer binding site hybridizes to the free 3′ end that is formed after a nick is formed in the non-target strand of the R-loop, thereby priming reverse transcriptase for DNA polymerization in the 5′-3′ direction.

[608] In one embodiment, the position of the intermolecular RNA extension is not in the protospacer sequence of the guide RNA. In another embodiment, the position of the intermolecular RNA extension in the gRNA core. In still another embodiment, the position of the intermolecular RNA extension is any with the guide RNA molecule except within the protospacer sequence, or at a position which disrupts the protospacer sequence.

In one embodiment, the intermolecular RNA extension is inserted downstream from the 3′ end of the protospacer sequence. In another embodiment, the intermolecular RNA extension is inserted at least 1 nucleotide, at least 2 nucleotides, at least 3 nucleotides, at least 4 nucleotides, at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13

nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 21 nucleotides, at least 22 nucleotides, at least 23 nucleotides, at least 24 nucleotides, at least 25 nucleotides downstream of the 3′ end of the protospacer sequence.

[609] In other embodiments, the intermolecular RNA extension is inserted into the gRNA, which refers to the portion of the guide RNA corresponding or comprising the tracrRNA, which binds and/or interacts with the Cas9 protein or equivalent thereof (i.e., a different napDNAbp). Preferably the insertion of the intermolecular RNA extension does not disrupt or minimally disrupts the interaction between the tracrRNA portion and the napDNAbp.

[610] The length of the RNA extension (which includes at least the RT template and primer binding site, e.g., see FIG. 3) can be any useful length. In various embodiments, the RNA extension is at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 21 nucleotides, at least 22 nucleotides, at least 23 nucleotides, at least 24 nucleotides, at least 25 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90 nucleotides, at least 100 nucleotides, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length.

[611] The RT template sequence can also be any suitable length. For example, the RT template sequence can be at least 3 nucleotides, at least 4 nucleotides, at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90 nucleotides, at least 100 nucleotides, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length.

[612] In still other embodiments, wherein the reverse transcription primer binding site sequence is at least 3 nucleotides, at least 4 nucleotides, at least 5 nucleotides, at least 6 nucleotides, at

least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90 nucleotides, at least 100 nucleotides, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length.

[613] In other embodiments, the optional linker or spacer sequence is at least 3 nucleotides, at least 4 nucleotides, at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90 nucleotides, at least 100 nucleotides, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length.

[614] The RT template sequence, in certain embodiments, encodes a single-stranded DNA molecule which is homologous to the non-target strand (and thus, complementary to the corresponding site of the target strand) but includes one or more nucleotide changes. The least one nucleotide change may include one or more single-base nucleotide changes, one or more deletions, and one or more insertions.

[615] The synthesized single-stranded DNA product of the RT template sequence is homologous to the non-target strand and contains one or more nucleotide changes. The single-stranded DNA product of the RT template sequence hybridizes in equilibrium with the complementary target strand sequence, thereby displacing the homologous endogenous target strand sequence. The displaced endogenous strand may be referred to in some embodiments as a 5′ endogenous DNA flap species. This 5′ endogenous DNA flap species can be removed by a 5′ flap endonuclease (e.g., FEN1) and the single-stranded DNA product, now hybridized to the endogenous target strand, may be ligated, thereby creating a mismatch between the endogenous sequence and the newly synthesized strand. The mismatch may be resolved by the cell's innate DNA repair and/or replication processes.

[616] In various embodiments, the nucleotide sequence of the RT template sequence corresponds to the nucleotide sequence of the non-target strand which becomes displaced as the 5′ flap species and which overlaps with the site to be edited.

[617] In various embodiments of the extended guide RNAs, the reverse transcription template sequence may encode a single-strand DNA flap that is complementary to an endogenous DNA sequence adjacent to a nick site, wherein the single-strand DNA flap comprises a desired nucleotide change. The single-stranded DNA flap may displace an endogenous single-strand DNA at the nick site. The displaced endogenous single-strand DNA at the nick site can have a 5′ end and form an endogenous flap, which can be excised by the cell. In various embodiments, excision of the 5′ end endogenous flap can help drive product formation since removing the 5′ end endogenous flap encourages hybridization of the single-strand 3′ DNA flap to the corresponding complementary DNA strand, and the incorporation or assimilation of the desired nucleotide change carried by the single-strand 3′ DNA flap into the target DNA.

[618] In various embodiments of the extended guide RNAs, the cellular repair of the single-strand DNA flap results in installation of the desired nucleotide change, thereby forming a desired product.

[619] In still other embodiments, the desired nucleotide change is installed in an editing window that is between about -5 to +5 of the nick site, or between about -10 to +10 of the nick site, or between about -20 to +20 of the nick site, or between about -30 to +30 of the nick site, or between about -40 to + 40 of the nick site, or between about -50 to +50 of the nick site, or between about -60 to +60 of the nick site, or between about -70 to +70 of the nick site, or between about -80 to +80 of the nick site, or between about -90 to +90 of the nick site, or between about -100 to +100 of the nick site, or between about -200 to +200 of the nick site.

[620] In other embodiments, the desired nucleotide change is installed in an editing window that is between about +1 to +2 from the nick site, or about +1 to +3, +1 to +4, +1 to +5, +1 to +6, +1 to +7, +1 to +8, +1 to +9, +1 to +10, +1 to +11, +1 to +12, +1 to +13, +1 to +14, +1 to +15, +1 to +16, +1 to +17, +1 to +18, +1 to +19, +1 to +20, +1 to +21, +1 to +22, +1 to +23, +1 to +24, +1 to +25, +1 to +26, +1 to +27, +1 to +28, +1 to +29, +1 to +30, +1 to +31, +1 to +32, +1 to +33, +1 to +34, +1 to +35, +1 to +36, +1 to +37, +1 to +38, +1 to +39, +1 to +40, +1 to +41, +1 to +42, +1 to +43, +1 to +44, +1 to +45, +1 to +46, +1 to +47, +1 to +48, +1 to +49, +1 to +50, +1 to +51, +1 to +52, +1 to +53, +1 to +54, +1 to +55, +1 to +56, +1 to +57, +1 to +58, +1

to +59, +1 to +60, +1 to +61, +1 to +62, +1 to +63, +1 to +64, +1 to +65, +1 to +66, +1 to +67, +1 to +68, +1 to +69, +1 to +70, +1 to +71, +1 to +72, +1 to +73, +1 to +74, +1 to +75, +1 to +76, +1 to +77, +1 to +78, +1 to +79, +1 to +80, +1 to +81, +1 to +82, +1 to +83, +1 to +84, +1 to +85, +1 to +86, +1 to +87, +1 to +88, +1 to +89, +1 to +90, +1 to +90, +1 to +91, +1 to +92, +1 to +93, +1 to +94, +1 to +95, +1 to +96, +1 to +97, +1 to +98, +1 to +99, +1 to +100, +1 to +101, +1 to +102, +1 to +103, +1 to +104, +1 to +105, +1 to +106, +1 to +107, +1 to +108, +1 to +109, +1 to +110, +1 to +111, +1 to +112, +1 to +113, +1 to +114, +1 to +115, +1 to +116, +1 to +117, +1 to +118, +1 to +119, +1 to +120, +1 to +121, +1 to +122, +1 to +123, +1 to +124, or +1 to +125 from the nick site.

[621] In still other embodiments, the desired nucleotide change is installed in an editing window that is between about +1 to +2 from the nick site, or about +1 to +5, +1 to +10, +1 to +15, +1 to +20, +1 to +25, +1 to +30, +1 to +35, +1 to +40, +1 to +45, +1 to +50, +1 to +55, +1 to +100, +1 to +105, +1 to +110, +1 to +115, +1 to +120, +1 to +125, +1 to +130, +1 to +135, +1 to +140, +1 to +145, +1 to +150, +1 to +155, +1 to +160, +1 to +165, +1 to +170, +1 to +175, +1 to +180, +1 to +185, +1 to +190, +1 to +195, or +1 to +200, from the nick site.

[622] In various aspects, the extended guide RNAs are modified versions of a guide RNA. Guide RNAs maybe naturally occurring, expressed from an encoding nucleic acid, or synthesized chemically. Methods are well known in the art for obtaining or otherwise synthesizing guide RNAs and for determining the appropriate sequence of the guide RNA, including the protospacer sequence which interacts and hybridizes with the target strand of a genomic target site of interest.

[623] In various embodiments, the particular design aspects of a guide RNA sequence will depend upon the nucleotide sequence of a genomic target site of interest (i.e., the desired site to be edited) and the type of napDNAbp (e.g., Cas9 protein) present in the prime editing systems utilized in the methods and compositions described herein, among other factors, such as PAM sequence locations, percent G/C content in the target sequence, the degree of microhomology regions, secondary structures, etc.

[624] In general, a guide sequence is any polynucleotide sequence having sufficient complementarity with a target polynucleotide sequence to hybridize with the target sequence and direct sequence-specific binding of a napDNAbp (e.g., a Cas9, Cas9 homolog, or Cas9 variant) to the target sequence. In some embodiments, the degree of complementarity between a guide

sequence and its corresponding target sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, non-limiting example of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (e.g., the Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, Calif.), SOAP (available at soap.genomics.org.cn), and Maq (available at maq.sourceforge.net). In some embodiments, a guide sequence is about or more than about 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75, or more nucleotides in length.

[625] In some embodiments, a guide sequence is less than about 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, or fewer nucleotides in length. The ability of a guide sequence to direct sequence-specific binding of a prime editor to a target sequence may be assessed by any suitable assay. For example, the components of a prime editor, including the guide sequence to be tested, may be provided to a host cell having the corresponding target sequence, such as by transfection with vectors encoding the components of a prime editor disclosed herein, followed by an assessment of preferential cleavage within the target sequence, such as by Surveyor assay as described herein. Similarly, cleavage of a target polynucleotide sequence may be evaluated in a test tube by providing the target sequence, components of a prime editor, including the guide sequence to be tested and a control guide sequence different from the test guide sequence, and comparing binding or rate of cleavage at the target sequence between the test and control guide sequence reactions. Other assays are possible, and will occur to those skilled in the art.

[626] A guide sequence may be selected to target any target sequence. In some embodiments, the target sequence is a sequence within a genome of a cell. Exemplary target sequences include those that are unique in the target genome. For example, for the S. pyogenes Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNXGG (SEQ ID NO: 173) where NNNNNNNNNNNNXGG (SEQ ID NO: 174) (N is A, G, T, or C; and X can be anything). A unique target sequence in a genome may include an S. pyogenes Cas9 target site of the form MMMMMMMMMNNNNNNNNNNNNXGG (SEQ ID NO: 175) where NNNNNNNNNNNXGG (SEQ ID NO: 176) (N is A, G, T, or C; and X can be anything). For the S. thermophilus

CRISPR1Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNXXAGAAW (SEQ ID NO: 177) where NNNNNNNNNNNNXXAGAAW (SEQ ID NO: 178) (N is A, G, T, or C; X can be anything; and W is A or T). A unique target sequence in a genome may include an S. thermophilus CRISPR 1 Cas9 target site of the form MMMMMMMMMNNNNNNNNNNNXXAGAAW (SEQ ID NO: 179) where NNNNNNNNNNNXXAGAAW (SEQ ID NO: 180) (N is A, G, T, or C; X can be anything; and W is A or T). For the S. pyogenes Cas9, a unique target sequence in a genome may include a Cas9 target site of the form MMMMMMMMNNNNNNNNNNNNXGGXG (SEQ ID NO: 181) where NNNNNNNNNNNNXGGXG (SEQ ID NO: 182) (N is A, G, T, or C; and X can be anything). A unique target sequence in a genome may include an S. pyogenes Cas9 target site of the form MMMMMMMMMNNNNNNNNNNNNXGGXG (SEQ ID NO: 183) where NNNNNNNNNNNNXGGXG (SEQ ID NO: 184) (N is A, G, T, or C; and X can be anything). In each of these sequences "M" may be A, G, T, or C, and need not be considered in identifying a sequence as unique.

[627] In some embodiments, a guide sequence is selected to reduce the degree of secondary structure within the guide sequence. Secondary structure may be determined by any suitable polynucleotide folding algorithm. Some programs are based on calculating the minimal Gibbs free energy. An example of one such algorithm is mFold, as described by Zuker and Stiegler (Nucleic Acids Res. 9 (1981), 133-148). Another example folding algorithm is the online webserver RNAfold, developed at Institute for Theoretical Chemistry at the University of Vienna, using the centroid structure prediction algorithm (see e.g., A. R. Gruber *et al.*, 2008, Cell 106(1): 23-24; and PA Carr and GM Church, 2009, Nature Biotechnology 27(12): 1151-62). Further algorithms may be found in U.S. application Ser. No. 61/836,080; Broad Reference BI-2013/004A); incorporated herein by reference.

[628] In general, a tracr mate sequence includes any sequence that has sufficient complementarity with a tracr sequence to promote one or more of: (1) excision of a guide sequence flanked by tracr mate sequences in a cell containing the corresponding tracr sequence; and (2) formation of a complex at a target sequence, wherein the complex comprises the tracr mate sequence hybridized to the tracr sequence. In general, degree of complementarity is with reference to the optimal alignment of the tracr mate sequence and tracr sequence, along the

length of the shorter of the two sequences. Optimal alignment may be determined by any suitable alignment algorithm, and may further account for secondary structures, such as self-complementarity within either the tracr sequence or tracr mate sequence. In some embodiments, the degree of complementarity between the tracr sequence and tracr mate sequence along the length of the shorter of the two when optimally aligned is about or more than about 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99%, or higher. In some embodiments, the tracr sequence is about or more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, or more nucleotides in length. In some embodiments, the tracr sequence and tracr mate sequence are contained within a single transcript, such that hybridization between the two produces a transcript having a secondary structure, such as a hairpin. Preferred loop forming sequences for use in hairpin structures are four nucleotides in length, and most preferably have the sequence GAAA. However, longer or shorter loop sequences may be used, as may alternative sequences. The sequences preferably include a nucleotide triplet (for example, AAA), and an additional nucleotide (for example C or G). Examples of loop forming sequences include CAAA and AAAG. In an embodiment of the invention, the transcript or transcribed polynucleotide sequence has at least two or more hairpins. In preferred embodiments, the transcript has two, three, four or five hairpins. In a further embodiment of the invention, the transcript has at most five hairpins. In some embodiments, the single transcript further includes a transcription termination sequence; preferably this is a polyT sequence, for example six T nucleotides. Further non-limiting examples of single polynucleotides comprising a guide sequence, a tracr mate sequence, and a tracr sequence are as follows (listed 5′ to 3′), where "N" represents a base of a guide sequence, the first block of lower case letters represent the tracr mate sequence, and the second block of lower case letters represent the tracr sequence, and the final poly-T sequence represents the transcription terminator:

(1)NNNNNNNNGTTTTTGTACTCTCAAGATTTAGAAATAAATCTTGCAGAAGCTACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTTCGTTATTTAATTTTTT (SEQ ID NO: 185);

(2)NNNNNNNNNNNNNNNNNNGTTTTTGTACTCTCAGAAATGCAGAAGCTACAAAGATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTTCGTTATTTAATTTTTT (SEQ ID NO: 186);

(3)NNNNNNNNNNNNNNNNNNNNGTTTTTGTACTCTCAGAAATGCAGAAGCTACAAA GATAAGGCTTCATGCCGAAATCAACACCCTGTCATTTTATGGCAGGGTGTTTTTT (SEQ ID NO: 187);
(4)NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 188);
(5)NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGG CTAGTCCGTTATCAACTTGAAAAAGTGTTTTTTT (SEQ ID NO: 189); and
(6)
NNNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT AGTCCGTTATCATTTTTTTT (SEQ ID NO: 190).

[629] In some embodiments, sequences (1) to (3) are used in combination with Cas9 from S. thermophilus CRISPR1. In some embodiments, sequences (4) to (6) are used in combination with Cas9 from S. pyogenes. In some embodiments, the tracr sequence is a separate transcript from a transcript comprising the tracr mate sequence.

[630] It will be apparent to those of skill in the art that in order to target any of the fusion proteins comprising a Cas9 domain and a single-stranded DNA binding protein, as disclosed herein, to a target site, e.g., a site comprising a point mutation to be edited, it is typically necessary to co-express the fusion protein together with a guide RNA, e.g., an sgRNA. As explained in more detail elsewhere herein, a guide RNA typically comprises a tracrRNA framework allowing for Cas9 binding, and a guide sequence, which confers sequence specificity to the Cas9:nucleic acid editing enzyme/domain fusion protein.

[631] In some embodiments, the guide RNA comprises a structure 5′-[guide sequence]-GUUUUAGAGCUAGAAAUAGCAAGUUAAAAUAAAGGCUAGUCCGUUAUCAACUUG AAAAAGUGGCACCGAGUCGGUGCUUUUU-3′ (SEQ ID NO: 191), wherein the guide sequence comprises a sequence that is complementary to the target sequence. The guide sequence is typically 20 nucleotides long. The sequences of suitable guide RNAs for targeting Cas9:nucleic acid editing enzyme/domain fusion proteins to specific genomic target sites will be apparent to those of skill in the art based on the instant disclosure. Such suitable guide RNA sequences typically comprise guide sequences that are complementary to a nucleic sequence within 50 nucleotides upstream or downstream of the target nucleotide to be edited. Some

exemplary guide RNA sequences suitable for targeting any of the provided fusion proteins to specific target sequences are provided herein. Additional guide sequences are well known in the art and can be used with the prime editors utilized in the methods and compositions described herein.

[632] In some embodiments, a PEgRNA comprises three main component elements ordered in the 5′ to 3′ direction, namely: a spacer, a gRNA core, and an extension arm at the 3′ end. The extension arm may further be divided into the following structural elements in the 5′ to 3′ direction, namely: a homology arm, an editing template, and a primer binding site. In addition, the PEgRNA may comprise an optional 3′ end modifier region (e1) and an optional 5′ end modifier region (e2). Still further, the PEgRNA may comprise a transcriptional termination signal at the 3′ end of the PEgRNA (not depicted). These structural elements are further defined herein. The depiction of the structure of the PEgRNA is not meant to be limiting and embraces variations in the arrangement of the elements. For example, the optional sequence modifiers (e1) and (e2) could be positioned within or between any of the other regions shown, and not limited to being located at the 3′ and 5′ ends.

[633] In some embodiments, a PEgRNA contemplated herein and may be designed in accordance with the methodology defined in Example 2. The PEgRNA comprises three main component elements ordered in the 5′ to 3′ direction, namely: a spacer, a gRNA core, and an extension arm at the 3′ end. The extension arm may further be divided into the following structural elements in the 5′ to 3′ direction, namely: a homology arm, an editing template, and a primer binding site. In addition, the PEgRNA may comprise an optional 3′ end modifier region (e1) and an optional 5′ end modifier region (e2). Still further, the PEgRNA may comprise a transcriptional termination signal on the 3′ end of the PEgRNA (not depicted). These structural elements are further defined herein. The depiction of the structure of the PEgRNA is not meant to be limiting and embraces variations in the arrangement of the elements. For example, the optional sequence modifiers (e1) and (e2) could be positioned within or between any of the other regions shown, and not limited to being located at the 3′ and 5′ ends.

PEgRNA improvements

[634] The PEgRNAs may also include additional design improvements that may modify the properties and/or characteristics of PEgRNAs thereby improving the efficacy of prime editing. In various embodiments, these improvements may belong to one or more of a number of different

categories, including but not limited to: (1) designs to enable efficient expression of functional PEgRNAs from non-polymerase III (pol III) promoters, which would enable the expression of longer PEgRNAs without burdensome sequence requirements; (2) improvements to the core, Cas9-binding PEgRNA scaffold, which could improve efficacy; (3) modifications to the PEgRNA to improve RT processivity, enabling the insertion of longer sequences at targeted genomic loci; and (4) addition of RNA motifs to the 5′ or 3′ termini of the PEgRNA that improve PEgRNA stability, enhance RT processivity, prevent misfolding of the PEgRNA, or recruit additional factors important for genome editing.

[635] In one embodiment, PEgRNA could be designed with polIII promoters to improve the expression of longer-length PEgRNA with larger extension arms. sgRNAs are typically expressed from the U6 snRNA promoter. This promoter recruits pol III to express the associated RNA and is useful for expression of short RNAs that are retained within the nucleus. However, pol III is not highly processive and is unable to express RNAs longer than a few hundred nucleotides in length at the levels required for efficient genome editing. Additionally, pol III can stall or terminate at stretches of U's, potentially limiting the sequence diversity that could be inserted using a PEgRNA. Other promoters that recruit polymerase II (such as pCMV) or polymerase I (such as the U1 snRNA promoter) have been examined for their ability to express longer sgRNAs. However, these promoters are typically partially transcribed, which would result in extra sequence 5′ of the spacer in the expressed PEgRNA, which has been shown to result in markedly reduced Cas9:sgRNA activity in a site-dependent manner. Additionally, while pol III-transcribed PEgRNAs can simply terminate in a run of 6-7 U's, PEgRNAs transcribed from pol II or pol I would require a different termination signal. Often such signals also result in polyadenylation, which would result in undesired transport of the PEgRNA from the nucleus. Similarly, RNAs expressed from pol II promoters such as pCMV are typically 5′-capped, also resulting in their nuclear export.

[636] Previously, Rinn and coworkers screened a variety of expression platforms for the production of long-noncoding RNA- (lncRNA) tagged sgRNAs[183]. These platforms include RNAs expressed from pCMV and that terminate in the ENE element from the MALAT1 ncRNA from humans[184], the PAN ENE element from KSHV[185], or the 3′ box from U1 snRNA[186]. Notably, the MALAT1 ncRNA and PAN ENEs form triple helices protecting the polyA-tail [184].

[187]. These constructs could also enhance RNA stability. It is contemplated that these expression systems will also enable the expression of longer PEgRNAs.

[637] In addition, a series of methods have been designed for the cleavage of the portion of the pol II promoter that would be transcribed as part of the PEgRNA, adding either a self-cleaving ribozyme such as the hammerhead[188], pistol[189], hatchet[189], hairpin[190], VS[191], twister[192], or twister sister[192] ribozymes, or other self-cleaving elements to process the transcribed guide, or a hairpin that is recognized by Csy4[193] and also leads to processing of the guide. Also, it is hypothesized that incorporation of multiple ENE motifs could lead to improved PEgRNA expression and stability, as previously demonstrated for the KSHV PAN RNA and element[185]. It is also anticipated that circularizing the PEgRNA in the form of a circular intronic RNA (ciRNA) could also lead to enhanced RNA expression and stability, as well as nuclear localization[194].

In various embodiments, the PEgRNA may include various above elements, as exemplified by the following sequence.

[638] Non-limiting example 1 - PEgRNA expression platform consisting of pCMV, Csy4 hairpin, the PEgRNA, and MALAT1 ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATA
ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGT
ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT
GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAA
ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGT
ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGA
CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG
AGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAG
ACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTTAGGGTCATGAAGGTTTTT
CTTTTCCTGAGAAAACAACACGTATTGTTTTCTCAGGTTTTGCTTTTTGGCCTTTTTCTAGCTTAAAAAA
AAAAAAAGCAAAAGATGCTGGTGGTTGGCACTCCTGGTTTCCAGGACGGGGTTCAAATCCCTGCGGCG
TCTTTGCTTTGACT (SEQ ID NO: 192)

[639] Non-limiting example 2 - PEgRNA expression platform consisting of pCMV, Csy4 hairpin, the PEgRNA, and PAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATA
ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGT
ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT
GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAA
ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGT
ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGA
CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG
AGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAG

ACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTTGTTTTGGCTGGGTTTTTC
CTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTTATCCCAGTGTATATTGGAAAAAC
ATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACT
TACAACATAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAA (SEQ ID NO: 193)

[640] Non-limiting example 3 - PEgRNA expression platform consisting of pCMV, Csy4 hairpin, the PEgRNA, and 3xPAN ENE

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATA
ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGT
ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT
GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAA
ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGT
ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGA
CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG
AGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAG
ACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTTGTTTTGGCTGGGTTTTTC
CTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTTATCCCAGTGTATATTGGAAAAAC
ATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACGTAATGCAACT
TACAACATAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAACACACTGTTTTGGC
TGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTTATCCCAGTGTATAT
TGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAATTAAACTAATACCATAACG
TAATGCAACTTACAACATAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAAAAATCTCTC
TGTTTTGGCTGGGTTTTTCCTTGTTCGCACCGGACACCTCCAGTGACCAGACGGCAAGGTTTTTTATCCC
AGTGTATATTGGAAAAACATGTTATACTTTTGACAATTTAACGTGCCTAGAGCTCAAATTAAACTAATA
CCATAACGTAATGCAACTTACAACATAAATAAAGGTCAATGTTTAATCCATAAAAAAAAAAAAAAAAA
AAA (SEQ ID NO: 194)

[641] Non-limiting example 4 - PEgRNA expression platform consisting of pCMV, Csy4 hairpin, the PEgRNA, and 3′ box

TAGTTATTAATAGTAATCAATTACGGGGTCATTAGTTCATAGCCCATATATGGAGTTCCGCGTTACATA
ACTTACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAATGACGT
ATGTTCCCATAGTAACGCCAATAGGGACTTTCCATTGACGTCAATGGGTGGAGTATTTACGGTAAACT
GCCCACTTGGCAGTACATCAAGTGTATCATATGCCAAGTACGCCCCCTATTGACGTCAATGACGGTAA
ATGGCCCGCCTGGCATTATGCCCAGTACATGACCTTATGGGACTTTCCTACTTGGCAGTACATCTACGT
ATTAGTCATCGCTATTACCATGGTGATGCGGTTTTGGCAGTACATCAATGGGCGTGGATAGCGGTTTGA
CTCACGGGGATTTCCAAGTCTCCACCCCATTGACGTCAATGGGAGTTTGTTTTGGCACCAAAATCAACG
GGACTTTCCAAAATGTCGTAACAACTCCGCCCCATTGACGCAAATGGGCGGTAGGCGTGTACGGTGGG
AGGTCTATATAAGCAGAGCTGGTTTAGTGAACCGTCAGATCGTTCACTGCCGTATAGGCAGGGCCCAG
ACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTGTTTCAAAAGTAGACTGT
ACGCTAAGGGTCATATCTTTTTTTGTTTGGTTTGTGTCTTGGTTGGCGTCTTAAA (SEQ ID NO: 195)

[642] Non-limiting example 5 - PEgRNA expression platform consisting of pU1, Csy4 hairpin, the PEgRNA, and 3′ box

CTAAGGACCAGCTTCTTTGGGAGAGAACAGACGCAGGGGCGGGAGGGAAAAAGGGAGAGGCAGACG
TCACTTCCCCTTGGCGGCTCTGGCAGCAGATTGGTCGGTTGAGTGGCAGAAAGGCAGACGGGGACTGG
GCAAGGCACTGTCGGTGACATCACGGACAGGGCGACTTCTATGTAGATGAGGCAGCGCAGAGGCTGC

TGCTTCGCCACTTGCTGCTTCACCACGAAGGAGTTCCCGTGCCCTGGGAGCGGGTTCAGGACCGCTGA
TCGGAAGTGAGAATCCCAGCTGTGTGTCAGGGCTGGAAAGGGCTCGGGAGTGCGCGGGGCAAGTGAC
CGTGTGTGTAAAGAGTGAGGCGTATGAGGCTGTGTCGGGGCAGAGGCCCAAGATCTCAGTTCACTGCC
GTATAGGCAGGGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTA
GTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTC
AGCAAGTTCAGAGAAATCTGAACTTGCTGGATTTTTGGAGCAGGGAGATGGAATAGGAGCTTGCTCCG
TCCACTCCACGCATCGACCTGGTATTGCAGTACCTCCAGGAACGGTGCACCCACTTTCTGGAGTTTCAA
AAGTAGACTGTACGCTAAGGGTCATATCTTTTTTTGTTTGGTTTGTGTCTTGGTTGGCGTCTTAAA (SEQ
ID NO: 196).

[643] In various other embodiments, the PEgRNA may be improved by introducing improvements to the scaffold or core sequences. This can be done by introducing known The core, Cas9-binding PEgRNA scaffold can likely be improved to enhance PE activity. Several such approaches have already been demonstrated. For instance, the first pairing element of the scaffold (P1) contains a GTTTT-AAAAC (SEQ ID NO: 68) pairing element. Such runs of Ts have been shown to result in pol III pausing and premature termination of the RNA transcript. Rational mutation of one of the T-A pairs to a G-C pair in this portion of P1 has been shown to enhance sgRNA activity, suggesting this approach would also be feasible for PEgRNAs[195]. Additionally, increasing the length of P1 has also been shown to enhance sgRNA folding and lead to improved activity[195], suggesting it as another avenue for the improvement of PEgRNA activity. Example improvements to the core can include:

PEgRNA containing a 6 nt extension to P1

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGCTCATGAAAATGAGCTAGCAAGTTAAAATAAGGCT
AGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTT
TTTT (SEQ ID NO: 197)

PEgRNA containing a T-A to G-C mutation within P1

GGCCCAGACTGAGCACGTGAGTTTGAGAGCTAGAAATAGCAAGTTTAAATAAGGCTAGTCCGTTATCA
ACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGTTTTTTT (SEQ ID
NO: 198)

[644] In various other embodiments, the PEgRNA may be improved by introducing modifications to the edit template region. As the size of the insertion templated by the PEgRNA increases, it is more likely to be degraded by endonucleases, undergo spontaneous hydrolysis, or fold into secondary structures unable to be reverse-transcribed by the RT or that disrupt folding of the PEgRNA scaffold and subsequent Cas9-RT binding. Accordingly, it is likely that modification to the template of the PEgRNA might be necessary to affect large insertions, such as the insertion of whole genes. Some strategies to do so include the incorporation of modified

nucleotides within a synthetic or semi-synthetic PEgRNA that render the RNA more resistant to degradation or hydrolysis or less likely to adopt inhibitory secondary structures[196]. Such modifications could include 8-aza-7-deazaguanosine, which would reduce RNA secondary structure in G-rich sequences; locked-nucleic acids (LNA) that reduce degradation and enhance certain kinds of RNA secondary structure; 2'-O-methyl, 2'-fluoro, or 2'-O-methoxyethoxy modifications that enhance RNA stability. Such modifications could also be included elsewhere in the PEgRNA to enhance stability and activity. Alternatively or additionally, the template of the PEgRNA could be designed such that it both encodes for a desired protein product and is also more likely to adopt simple secondary structures that are able to be unfolded by the RT. Such simple structures would act as a thermodynamic sink, making it less likely that more complicated structures that would prevent reverse transcription would occur. Finally, one could also split the template into two, separate PEgRNAs. In such a design, a PE would be used to initiate transcription and also recruit a separate template RNA to the targeted site via an RNA-binding protein fused to Cas9 or an RNA recognition element on the PEgRNA itself such as the MS2 aptamer. The RT could either directly bind to this separate template RNA, or initiate reverse transcription on the original PEgRNA before swapping to the second template. Such an approach could enable long insertions by both preventing misfolding of the PEgRNA upon addition of the long template and also by not requiring dissociation of Cas9 from the genome for long insertions to occur, which could possibly be inhibiting PE-based long insertions.

[645] In still other embodiments, the PEgRNA may be improved by introducing additional RNA motifs at the 5′ and 3′ termini of the PEgRNAs, or even at positions therein between (e.g., in the gRNA core region, or the spacer). Several such motifs - such as the PAN ENE from KSHV and the ENE from MALAT1 were discussed above as possible means to terminate expression of longer PEgRNAs from non-pol III promoters. These elements form RNA triple helices that engulf the polyA tail, resulting in their being retained within the nucleus[184, 187]. However, by forming complex structures at the 3′ terminus of the PEgRNA that occlude the terminal nucleotide, these structures would also likely help prevent exonuclease-mediated degradation of PEgRNAs.

[646] Other structural elements inserted at the 3′ terminus could also enhance RNA stability, albeit without enabling termination from non-pol III promoters. Such motifs could include hairpins or RNA quadruplexes that would occlude the 3′ terminus[197], or self-cleaving ribozymes

such as HDV that would result in the formation of a 2'-3'-cyclic phosphate at the 3' terminus and also potentially render the PEgRNA less likely to be degraded by exonucleases[198]. Inducing the PEgRNA to cyclize via incomplete splicing - to form a ciRNA - could also increase PEgRNA stability and result in the PEgRNA being retained within the nucleus[194].

Additional RNA motifs could also improve RT processivity or enhance PEgRNA activity by enhancing RT binding to the DNA-RNA duplex. Addition of the native sequence bound by the RT in its cognate retroviral genome could enhance RT activity[199]. This could include the native primer binding site (PBS), polypurine tract (PPT), or kissing loops involved in retroviral genome dimerization and initiation of transcription[199].

[647] Addition of dimerization motifs - such as kissing loops or a GNRA tetraloop/tetraloop receptor pair[200] - at the 5′ and 3′ termini of the PEgRNA could also result in effective circularization of the PEgRNA, improving stability. Additionally, it is envisioned that addition of these motifs could enable the physical separation of the PEgRNA spacer and primer binding site, preventing occlusion of the spacer which would hinder PE activity. Short 5′ extensions or 3′ extensions to the PEgRNA that form a small toehold hairpin in the spacer region or along the primer binding site could also compete favorably against the annealing of intracomplementary regions along the length of the PEgRNA, e.g., the interaction between the spacer and the primer binding site that can occur. Finally, kissing loops could also be used to recruit other template RNAs to the genomic site and enable swapping of RT activity from one RNA to the other. A number secondary RNA structures that may be engineered into any region of the PEgRNA, including in the terminal portions of the extension arm (i.e., e1and e2), as shown.

Example improvements include, but are not limited to:

[648] PEgRNA-HDV fusion

GGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA ACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCGTGCTCAGTCTGGGCCGGCATGGTCC CAGCCTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTTTTTT (SEQ ID NO: 199)

[649] PEgRNA-MMLV kissing loop

GGTGGGAGACGTCCCACCGGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCTTCGACC GTGCTCAGTCTGGTGGGAGACGTCCCACCTTTTTTT (SEQ ID NO: 200)

[650] PEgRNA-VS ribozyme kissing loop

GAGCAGCATGGCGTCGCTGCTCACGGCCCAGACTGAGCACGTGAGTTTTAGAGCTAGAAATAGCAAGT
TAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCCTCTGCCATCAAAGCTT
CGACCGTGCTCAGTCTCCATCAGTTGACACCCTGAGGTTTTTTT (SEQ ID NO: 201)

[651] PEgRNA-GNRA tetraloop/tetraloop receptor

GCAGACCTAAGTGGUGACATATGGTCTGGGCCCAGACTGAGCACGTGAGTTTTAGAGCTAUACGTAGC
AAGTTAAAATAAGGCTAGTCCGTTATCAACTTUACGAAGTGGGACCGAGTCGGTCCTCTGCCATCAAA
GCTTCGACCGTGCTCAGTCTGCATGCGATTAGAAATAATCGCATGTTTTTTT (SEQ ID NO: 202)

[652] PEgRNA template switching secondary RNA-HDV fusion

TCTGCCATCAAAGCTGCGACCGTGCTCAGTCTGGTGGGAGACGTCCCACCGGCCGGCATGGTCCCAGC
CTCCTCGCTGGCGCCGGCTGGGCAACATGCTTCGGCATGGCGAATGGGACTTTTTTT (SEQ ID NO: 203)

[653] PEgRNA scaffolds could be further improved via directed evolution, in an analogous fashion to how SpCas9 and prime editors (PE) have been improved. Directed evolution could enhance PEgRNA recognition by Cas9 or evolved Cas9 variants. Additionally, it is likely that different PEgRNA scaffold sequences would be optimal at different genomic loci, either enhancing PE activity at the site in question, reducing off-target activities, or both. Finally, evolution of PEgRNA scaffolds to which other RNA motifs have been added would almost certainly improve the activity of the fused PEgRNA relative to the unevolved, fusion RNA. For instance, evolution of allosteric ribozymes composed of c-di-GMP-I aptamers and hammerhead ribozymes led to dramatically improved activity[202], suggesting that evolution would improve the activity of hammerhead-PEgRNA fusions as well. In addition, while Cas9 currently does not generally tolerate 5′ extension of the sgRNA, directed evolution will likely generate enabling mutations that mitigate this intolerance, allowing additional RNA motifs to be utilized.
The present disclosure contemplates any such ways to further improve the efficacy of the prime editing systems utilized in the methods and compositions disclosed here.

[654] In various embodiments, it may be advantageous to limit the appearance of consecutive sequence of Ts from the extension arm as consecutive series of T's may limit the capacity of the PEgRNA to be transcribed. For example, strings of at least consecutive three T's, at least consecutive four T's, at least consecutive five T's, at least consecutive six T's, at least consecutive seven T's, at least consecutive eight T's, at least consecutive nine T's, at least consecutive ten T's, at least consecutive eleven T's, at least consecutive twelve T's, at least consecutive thirteen T's , at least consecutive fourteen T's, or at least consecutive fifteen T's should be avoided when designing the PEgRNA, or should be at least removed from the final designed sequence. In one embodiment, one can avoid the includes of unwanted strings of

consecutive T's in PEgRNA extension arms but avoiding target sites that are rich in consecutive A:T nucleobase pairs.

PEgRNAs for evading MMR

[655] The present disclosure also provides novel pegRNAs for use in prime editing. As discussed above, prime editing using pegRNAs having DNA synthesis templates that comprise three or more consecutive nucleotide mismatches relative to a target site sequence can evade correction by the MMR pathway, resulting in an increase in prime editing efficiency and/or a decrease in the frequency of indel formation compared to the introduction of a single nucleotide mismatch using prime editing. Thus, the present disclosure provides pegRNAs useful for introducing modifications into a target nucleic acid with increased prime editing efficiency and/or decreased indel frequency compared to a corresponding control pegRNA that does not contain three or more consecutive nucleotide mismatches relative to the target site sequence..

[656] The pegRNAs provided by the present disclosure are useful for editing a nucleic acid molecule by prime editing while improving prime editing efficiency and/or reducing indel formation. Without wishing to be bound by theory, the pegRNAs provided in the present disclosure may evade or reduce the impact of cellular MMR correction of mismatches at the target site that are introduced by the nucleotide alteration(s) through prime editing. In some embodiments, the extension arm of the pegRNAs provided by the present disclosure comprise three or more consecutive nucleotide mismatches relative to a target site on the nucleic acid molecule. In some embodiments, the DNA synthesis template of the pegRNA comprises three or more consecutive nucleotide mismatches relative to the target site on the nucleic acid molecule. In some embodiments, at least one of the three or more consecutive nucleotide mismatches introduces a silent mutation. In some embodiments, at least one of the consecutive nucleotide mismatches results in an alteration in the amino acid sequence of a protein expressed from the target nucleic acid molecule, while at least one of the remaining nucleotide mismatches is a silent mutation. The silent mutations may be introduced in coding regions of the target nucleic acid molecule or in non-coding regions of the target nucleic acid molecule. When the silent mutations are introduced in a coding region, they may introduce into the nucleic acid molecule one or more alternate codons encoding the same amino acid as the unedited nucleic acid molecule. When the silent mutations are introduced in a non-coding region, they may be present in a region of the

nucleic acid molecule that does not influence splicing, gene regulation, RNA lifetime, or other biological properties of the target site on the nucleic acid molecule.

[657] Any number of consecutive nucleotide mismatches of three or more can be incorporated into the extension arm of the pegRNAs described herein to achieve the benefits of avoiding or reducing the impact of correction by the cellular MMR pathway. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises three or more, four or more, five or more, six or more, seven or more, eight or more, nine or more, or ten or more consecutive nucleotide mismatches relative to a target site on the nucleic acid molecule.

[658] Any number of consecutive nucleotide mismatches of three or more can be incorporated into the extension arm of the pegRNAs described herein to achieve the benefits of evading correction by the MMR pathway, thereby increasing editing efficiency and/or reducing unintended indels. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises three consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 3-5 consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 6-10 consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 11-20 consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 20-25 consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing. In some embodiments, the DNA synthesis template of the extension arm on the pegRNA comprises three or more, four or more, five or more, six or more, seven or more, eight or more, nine or

more, or ten or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule edited by prime editing.

Kits

[659] The compositions of the present disclosure may be assembled into kits. In some embodiments, the kit comprises nucleic acid vectors for the expression of a prime editor and an MMR inhibitor, such as, but not limited to an MLH1 dominant negative variant as described herein. In other embodiments, the kit further comprises appropriate guide nucleotide sequences (*e.g.*, PEgRNAs and second-site gRNAs) or nucleic acid vectors for the expression of such guide nucleotide sequences, to target the Cas9 protein or prime editor to the desired target sequence.

[660] The kit described herein may include one or more containers housing components for performing the methods described herein and optionally instructions for use. Any of the kit described herein may further comprise components needed for performing the assay methods. Each component of the kits, where applicable, may be provided in liquid form (*e.g.*, in solution) or in solid form, (*e.g.*, a dry powder). In certain cases, some of the components may be reconstitutable or otherwise processible (*e.g.*, to an active form), for example, by the addition of a suitable solvent or other species (for example, water), which may or may not be provided with the kit.

[661] In some embodiments, the kits may optionally include instructions and/or promotion for use of the components provided. As used herein, "instructions" can define a component of instruction and/or promotion, and typically involve written instructions on or associated with packaging of the disclosure. Instructions also can include any oral or electronic instructions provided in any manner such that a user will clearly recognize that the instructions are to be associated with the kit, for example, audiovisual (*e.g.*, videotape, DVD, *etc.*), Internet, and/or web-based communications, *etc.* The written instructions may be in a form prescribed by a governmental agency regulating the manufacture, use, or sale of pharmaceuticals or biological products, which can also reflect approval by the agency of manufacture, use or sale for animal administration. As used herein, "promoted" includes all methods of doing business including methods of education, hospital and other clinical instruction, scientific inquiry, drug discovery or development, academic research, pharmaceutical industry activity including pharmaceutical sales, and any advertising or other promotional activity including written, oral and electronic

communication of any form, associated with the disclosure. Additionally, the kits may include other components depending on the specific application, as described herein.

[662] The kits may contain any one or more of the components described herein in one or more containers. The components may be prepared sterilely, packaged in a syringe and shipped refrigerated. Alternatively, it may be housed in a vial or other container for storage. A second container may have other components prepared sterilely. Alternatively, the kits may include the active agents premixed and shipped in a vial, tube, or other container.

[663] The kits may have a variety of forms, such as a blister pouch, a shrink wrapped pouch, a vacuum sealable pouch, a sealable thermoformed tray, or a similar pouch or tray form, with the accessories loosely packed within the pouch, one or more tubes, containers, a box or a bag. The kits may be sterilized after the accessories are added, thereby allowing the individual accessories in the container to be otherwise unwrapped. The kits can be sterilized using any appropriate sterilization techniques, such as radiation sterilization, heat sterilization, or other sterilization methods known in the art. The kits may also include other components, depending on the specific application, for example, containers, cell media, salts, buffers, reagents, syringes, needles, a fabric, such as gauze, for applying or removing a disinfecting agent, disposable gloves, a support for the agents prior to administration, *etc*. Some aspects of this disclosure provide kits comprising a nucleic acid construct comprising a nucleotide sequence encoding the various components of the prime editing system utilized in the methods and compositions described herein (e.g., including, but not limited to, the napDNAbps, reverse transcriptases, polymerases, fusion proteins (e.g., comprising napDNAbps and reverse transcriptases (or more broadly, polymerases), extended guide RNAs, and complexes comprising fusion proteins and extended guide RNAs, as well as accessory elements, such as second strand nicking components (e.g., second strand nicking gRNA) and 5′ endogenous DNA flap removal endonucleases for helping to drive the prime editing process towards the edited product formation). In some embodiments, the nucleotide sequence(s) comprises a heterologous promoter (or more than a single promoter) that drives expression of the prime editing system components.

[664] Other aspects of this disclosure provide kits comprising one or more nucleic acid constructs encoding the various components of the prime editing systems utilized in the methods and compositions described herein, e.g., the comprising a nucleotide sequence encoding the components of the prime editing system capable of modifying a target DNA sequence. In some

embodiments, the nucleotide sequence comprises a heterologous promoter that drives expression of the prime editing system components.

[665] Some aspects of this disclosure provides kits comprising a nucleic acid construct, comprising (a) a nucleotide sequence encoding a napDNAbp (*e.g.*, a Cas9 domain) fused to a reverse transcriptase and (b) a heterologous promoter that drives expression of the sequence of (a).

Cells

[666] Cells that may contain any of the compositions described herein include prokaryotic cells and eukaryotic cells. The methods described herein are used to deliver a Cas9 protein or a prime editor and an MMR inhibitor (e.g., an MLH1 dominant negative variant) into a eukaryotic cell (*e.g.*, a mammalian cell, such as a human cell). In some embodiments, the cell is *in vitro* (*e.g.*, cultured cell. In some embodiments, the cell is *in vivo* (*e.g.*, in a subject such as a human subject). In some embodiments, the cell is *ex vivo* (*e.g.*, isolated from a subject and may be administered back to the same or a different subject).

Mammalian cells of the present disclosure include human cells, primate cells (*e.g.*, vero cells), rat cells (*e.g.*, GH3 cells, OC23 cells) or mouse cells (*e.g.*, MC3T3 cells). There are a variety of human cell lines, including, without limitation, human embryonic kidney (HEK) cells, HeLa cells, cancer cells from the National Cancer Institute's 60 cancer cell lines (NCI60), DU145 (prostate cancer) cells, Lncap (prostate cancer) cells, MCF-7 (breast cancer) cells, MDA-MB-438 (breast cancer) cells, PC3 (prostate cancer) cells, T47D (breast cancer) cells, THP-1 (acute myeloid leukemia) cells, U87 (glioblastoma) cells, SHSY5Y human neuroblastoma cells (cloned from a myeloma) and Saos-2 (bone cancer) cells. In some embodiments, rAAV vectors are delivered into human embryonic kidney (HEK) cells (*e.g.*, HEK 293 or HEK 293T cells). In some embodiments, rAAV vectors are delivered into stem cells (*e.g.*, human stem cells) such as, for example, pluripotent stem cells (*e.g.*, human pluripotent stem cells including human induced pluripotent stem cells (hiPSCs)). A stem cell refers to a cell with the ability to divide for indefinite periods in culture and to give rise to specialized cells. A pluripotent stem cell refers to a type of stem cell that is capable of differentiating into all tissues of an organism, but not alone capable of sustaining full organismal development. A human induced pluripotent stem cell refers to a somatic (*e.g.*, mature or adult) cell that has been reprogrammed to an embryonic stem cell-like state by being forced to express genes and factors important for maintaining the defining

properties of embryonic stem cells (*see*, *e.g.*, Takahashi and Yamanaka, *Cell* 126 (4): 663–76, 2006, incorporated by reference herein). Human induced pluripotent stem cell cells express stem cell markers and are capable of generating cells characteristic of all three germ layers (ectoderm, endoderm, mesoderm).

[667] Additional non-limiting examples of cell lines that may be used in accordance with the present disclosure include 293-T, 293-T, 3T3, 4T1, 721, 9L, A-549, A172, A20, A253, A2780, A2780ADR, A2780cis, A431, ALC, B16, B35, BCP-1, BEAS-2B, bEnd.3, BHK-21, BR 293, BxPC3, C2C12, C3H-10T1/2, C6, C6/36, Cal-27, CGR8, CHO, CML T1, CMT, COR-L23, COR-L23/5010, COR-L23/CPR, COR-L23/R23, COS-7, COV-434, CT26, D17, DH82, DU145, DuCaP, E14Tg2a, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, Hepa1c1c7, High Five cells, HL-60, HMEC, HT-29, HUVEC, J558L cells, Jurkat, JY cells, K562 cells, KCL22, KG1, Ku812, KYO1, LNCap, Ma-Mel 1, 2, 3. . 48, MC-38, MCF-10A, MCF-7, MDA-MB-231, MDA-MB-435, MDA-MB-468, MDCK II, MG63, MONO-MAC 6, MOR/0.2R, MRC5, MTD-1A, MyEnd, NALM-1, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NW-145, OPCN/OPCT Peer, PNT-1A/PNT 2, PTK2, Raji, RBL cells, RenCa, RIN-5F, RMA/RMAS, S2, Saos-2 cells, Sf21, Sf9, SiHa, SKBR3, SKOV-3, T-47D, T2, T84, THP1, U373, U87, U937, VCaP, WM39, WT-49, X63, YAC-1 and YAR cells.

[668] Some aspects of this disclosure provide cells comprising any of the constructs disclosed herein. In some embodiments, a host cell is transiently or non-transiently transfected with one or more vectors described herein. In some embodiments, a cell is transfected ex vivo. In some embodiments, a cell is transfected in vivo. In some embodiments, a cell that is transfected is taken from a subject. In some embodiments, the cell is derived from cells taken from a subject, such as a cell line. A wide variety of cell lines for tissue culture are known in the art. Examples of cell lines include, but are not limited to, C8161, CCRF-CEM, MOLT, mIMCD-3, NHDF, HeLa-S3, Huh1, Huh4, Huh7, HUVEC, HASMC, HEKn, HEKa, MiaPaCell, Panc1, PC-3, TF1, CTLL-2, C1R, Rat6, CV1, RPTE, A10, T24, J82, A375, ARH-77, Calu1, SW480, SW620, SKOV3, SK-UT, CaCo2, P388D1, SEM-K2, WEHI-231, HB56, TIB55, Jurkat, J45.01, LRMB, Bcl-1, BC-3, IC21, DLD2, Raw264.7, NRK, NRK-52E, MRC5, MEF, Hep G2, HeLa B, HeLa T4, COS, COS-1, COS-6, COS-M6A, BS-C-1 monkey kidney epithelial, BALB/3T3 mouse embryo fibroblast, 3T3 Swiss, 3T3-L1, 132-d5 human fetal fibroblasts; 10.1 mouse fibroblasts,

293-T, 3T3, 721, 9L, A2780, A2780ADR, A2780cis, A 172, A20, A253, A431, A-549, ALC, B16, B35, BCP-1 cells, BEAS-2B, bEnd.3, BHK-21, BR 293. BxPC3. C3H-10T1/2, C6/36, Cal-27, CHO, CHO-7, CHO-IR, CHO-K1, CHO-K2, CHO-T, CHO Dhfr −/−, COR-L23, COR-L23/CPR, COR-L23/5010, COR-L23/R23, COS-7, COV-434, CML T1, CMT, CT26, D17, DH82, DU145, DuCaP, EL4, EM2, EM3, EMT6/AR1, EMT6/AR10.0, FM3, H1299, H69, HB54, HB55, HCA2, HEK-293, HeLa, Hepa1c1c7, HL-60, HMEC, HT-29, Jurkat, JY cells, K562 cells, Ku812, KCL22, KG1, KYO1, LNCap, Ma-Mel 1-48, MC-38, MCF-7, MCF-10A, MDA-MB-231, MDA-MB-468, MDA-MB-435, MDCK II, MDCK 11, MOR/0.2R, MONO-MAC 6, MTD-1A, MyEnd, NCI-H69/CPR, NCI-H69/LX10, NCI-H69/LX20, NCI-H69/LX4, NIH-3T3, NALM-1, NW-145, OPCN/OPCT cell lines, Peer, PNT-1A/PNT 2, RenCa, RIN-5F, RMA/RMAS, Saos-2 cells, Sf-9, SkBr3, T2, T-47D, T84, THP1 cell line, U373, U87, U937, VCaP, Vero cells, WM39, WT-49, X63, YAC-1, YAR, and transgenic varieties thereof.

[669] Cell lines are available from a variety of sources known to those with skill in the art (see, e.g., the American Type Culture Collection (ATCC) (Manassus, Va.)). In some embodiments, a cell transfected with one or more vectors described herein is used to establish a new cell line comprising one or more vector-derived sequences. In some embodiments, a cell transiently transfected with the components of a CRISPR system as described herein (such as by transient transfection of one or more vectors, or transfection with RNA), and modified through the activity of a CRISPR complex, is used to establish a new cell line comprising cells containing the modification but lacking any other exogenous sequence. In some embodiments, cells transiently or non-transiently transfected with one or more vectors described herein, or cell lines derived from such cells are used in assessing one or more test compounds.

Vectors

[670] Some aspects of the present disclosure relate to using recombinant virus vectors (e.g., adeno-associated virus vectors, adenovirus vectors, or herpes simplex virus vectors) for the delivery of the prime editors and MLH1 dominant negative mutants as described herein into a cell. In the case of a split-PE approach, the N-terminal portion of a PE fusion protein and the C-terminal portion of a PE fusion are delivered by separate recombinant virus vectors (e.g., adeno-associated virus vectors, adenovirus vectors, or herpes simplex virus vectors) into the same cell, since the full-length Cas9 protein or prime editors exceeds the packaging limit of various virus vectors, e.g., rAAV (~4.9 kb).

[671] In some embodiments, the vectors used herein may encode the PE fusion proteins, or any of the components thereof (e.g., napDNAbp, linkers, or polymerases), or an MLH1 dominant negative mutant. In addition, the vectors used herein may encode the PEgRNAs, and/or the accessory gRNA for second strand nicking. The vectors may be capable of driving expression of one or more coding sequences in a cell. In some embodiments, the cell may be a prokaryotic cell, such as, e.g., a bacterial cell. In some embodiments, the cell may be a eukaryotic cell, such as, e.g., a yeast, plant, insect, or mammalian cell. In some embodiments, the eukaryotic cell may be a mammalian cell. In some embodiments, the eukaryotic cell may be a rodent cell. In some embodiments, the eukaryotic cell may be a human cell. Suitable promoters to drive expression in different types of cells are known in the art. In some embodiments, the promoter may be wild-type. In other embodiments, the promoter may be modified for more efficient or efficacious expression. In yet other embodiments, the promoter may be truncated yet retain its function. For example, the promoter may have a normal size or a reduced size that is suitable for proper packaging of the vector into a virus.

[672] In some embodiments, the promoters that may be used in the prime editor vectors may be constitutive, inducible, or tissue-specific. In some embodiments, the promoters may be a constitutive promoters. Non-limiting exemplary constitutive promoters include cytomegalovirus immediate early promoter (CMV), simian virus (SV40) promoter, adenovirus major late (MLP) promoter, Rous sarcoma virus (RSV) promoter, mouse mammary tumor virus (MMTV) promoter, phosphoglycerate kinase (PGK) promoter, elongation factor-alpha (EFla) promoter, ubiquitin promoters, actin promoters, tubulin promoters, immunoglobulin promoters, a functional fragment thereof, or a combination of any of the foregoing. In some embodiments, the promoter may be a CMV promoter. In some embodiments, the promoter may be a truncated CMV promoter. In other embodiments, the promoter may be an EFla promoter. In some embodiments, the promoter may be an inducible promoter. Non-limiting exemplary inducible promoters include those inducible by heat shock, light, chemicals, peptides, metals, steroids, antibiotics, or alcohol. In some embodiments, the inducible promoter may be one that has a low basal (non-induced) expression level, such as, e.g., the Tet-On® promoter (Clontech). In some embodiments, the promoter may be a tissue-specific promoter. In some embodiments, the tissue-specific promoter is exclusively or predominantly expressed in liver tissue. Non-limiting exemplary tissue-specific promoters include B29 promoter, CD14 promoter, CD43 promoter,

CD45 promoter, CD68 promoter, desmin promoter, elastase- 1 promoter, endoglin promoter, fibronectin promoter, Flt-1 promoter, GFAP promoter, GPIIb promoter, ICAM- 2 promoter, INF-β promoter, Mb promoter, Nphsl promoter, OG-2 promoter, SP-B promoter, SYN1 promoter, and WASP promoter.

[673] In some embodiments, the prime editor and MLH1 dominant negative mutant vectors (e.g., including any vectors encoding the prime editor fusion protein and/or the PEgRNAs, and/or the accessory second strand nicking gRNAs) may comprise inducible promoters to start expression only after it is delivered to a target cell. Non-limiting exemplary inducible promoters include those inducible by heat shock, light, chemicals, peptides, metals, steroids, antibiotics, or alcohol. In some embodiments, the inducible promoter may be one that has a low basal (non-induced) expression level, such as, e.g., the Tet-On® promoter (Clontech).

[674] In additional embodiments, the prime editor vectors (e.g., including any vectors encoding the prime editor fusion protein and/or the PEgRNAs, and/or the accessory second strand nicking gRNAs) and MLH1 dominant negative mutant vector (e.g., any vector encoding an MLH1 dominant negative mutant as described herein) may comprise tissue- specific promoters to start expression only after it is delivered into a specific tissue. Non-limiting exemplary tissue-specific promoters include B29 promoter, CD14 promoter, CD43 promoter, CD45 promoter, CD68 promoter, desmin promoter, elastase- 1 promoter, endoglin promoter, fibronectin promoter, Flt-1 promoter, GFAP promoter, GPIIb promoter, ICAM- 2 promoter, INF-β promoter, Mb promoter, Nphsl promoter, OG-2 promoter, SP-B promoter, SYN1 promoter, and WASP promoter.

[675] In some embodiments, the nucleotide sequence encoding the PEgRNA (or any guide RNAs used in connection with prime editing) may be operably linked to at least one transcriptional or translational control sequence. In some embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to at least one promoter. In some embodiments, the promoter may be recognized by RNA polymerase III (Pol III). Non-limiting examples of Pol III promoters include U6, HI and tRNA promoters. In some embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to a mouse or human U6 promoter. In other embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to a mouse or human HI promoter. In some embodiments, the nucleotide sequence encoding the guide RNA may be operably linked to a mouse or human tRNA promoter. In embodiments with more than one guide RNA, the promoters used to drive expression may be the same or different.

In some embodiments, the nucleotide encoding the crRNA of the guide RNA and the nucleotide encoding the tracr RNA of the guide RNA may be provided on the same vector. In some embodiments, the nucleotide encoding the crRNA and the nucleotide encoding the tracr RNA may be driven by the same promoter. In some embodiments, the crRNA and tracr RNA may be transcribed into a single transcript. For example, the crRNA and tracr RNA may be processed from the single transcript to form a double-molecule guide RNA. Alternatively, the crRNA and tracr RNA may be transcribed into a single-molecule guide RNA.

[676] In some embodiments, the nucleotide sequence encoding the guide RNA may be located on the same vector comprising the nucleotide sequence encoding the PE fusion protein. In some embodiments, expression of the guide RNA and of the PE fusion protein may be driven by their corresponding promoters. In some embodiments, expression of the guide RNA may be driven by the same promoter that drives expression of the PE fusion protein. In some embodiments, the guide RNA and the PE fusion protein transcript may be contained within a single transcript. For example, the guide RNA may be within an untranslated region (UTR) of the Cas9 protein transcript. In some embodiments, the guide RNA may be within the 5′ UTR of the PE fusion protein transcript. In other embodiments, the guide RNA may be within the 3' UTR of the PE fusion protein transcript. In some embodiments, the intracellular half-life of the PE fusion protein transcript may be reduced by containing the guide RNA within its 3' UTR and thereby shortening the length of its 3' UTR. In additional embodiments, the guide RNA may be within an intron of the PE fusion protein transcript. In some embodiments, suitable splice sites may be added at the intron within which the guide RNA is located such that the guide RNA is properly spliced out of the transcript. In some embodiments, expression of the Cas9 protein and the guide RNA in close proximity on the same vector may facilitate more efficient formation of the CRISPR complex.

[677] The vector system may comprise one vector, or two vectors, or three vectors, or four vectors, or five vector, or more. In some embodiments, the vector system may comprise one single vector, which encodes both the PE fusion protein, the PEgRNA, and an MLH1 dominant negative mutant. In other embodiments, the vector system may comprise two vectors, wherein one vector encodes the PE fusion protein and the PEgRNA, and the other encodes the MLH1 dominant negative mutant.

[678] Some examples of materials which can serve as pharmaceutically-acceptable carriers include: (1) sugars, such as lactose, glucose and sucrose; (2) starches, such as corn starch and potato starch; (3) cellulose, and its derivatives, such as sodium carboxymethyl cellulose, methylcellulose, ethyl cellulose, microcrystalline cellulose and cellulose acetate; (4) powdered tragacanth; (5) malt; (6) gelatin; (7) lubricating agents, such as magnesium stearate, sodium lauryl sulfate and talc; (8) excipients, such as cocoa butter and suppository waxes; (9) oils, such as peanut oil, cottonseed oil, safflower oil, sesame oil, olive oil, corn oil and soybean oil; (10) glycols, such as propylene glycol; (11) polyols, such as glycerin, sorbitol, mannitol and polyethylene glycol (PEG); (12) esters, such as ethyl oleate and ethyl laurate; (13) agar; (14) buffering agents, such as magnesium hydroxide and aluminum hydroxide; (15) alginic acid; (16) pyrogen-free water; (17) isotonic saline; (18) Ringer's solution; (19) ethyl alcohol; (20) pH buffered solutions; (21) polyesters, polycarbonates and/or polyanhydrides; (22) bulking agents, such as polypeptides and amino acids (23) serum component, such as serum albumin, HDL and LDL; (22) C2-C12 alcohols, such as ethanol; and (23) other non-toxic compatible substances employed in pharmaceutical formulations. Wetting agents, coloring agents, release agents, coating agents, sweetening agents, flavoring agents, perfuming agents, preservative and antioxidants can also be present in the formulation. The terms such as "excipient", "carrier", "pharmaceutically acceptable carrier" or the like are used interchangeably herein.

Delivery methods

[679] In some aspects, the invention provides methods comprising delivering one or more polynucleotides, such as or one or more vectors as described herein, one or more transcripts thereof, and/or one or proteins transcribed therefrom, to a host cell. In some aspects, the invention further provides cells produced by such methods, and organisms (such as animals, plants, or fungi) comprising or produced from such cells. In some embodiments, a prime editor as described herein in combination with (and optionally complexed with) a guide sequence, as well as an inhibitor of the DNA mismatch repair pathway, are delivered to a cell. In any of the delivery methods described herein, an inhibitor of the DNA mismatch repair pathway can also be delivered along with the prime editor. In some embodiments, the inhibitor is MLH1dn as described further herein. In some embodiments, the inhibitor is encoded on the same vector as the prime editor. In certain embodiments, the inhibitor is fused to the prime editor. In some embodiments, the inhibitor is encoded on a second vector, which is delivered along with a vector

encoding the prime editor. In some embodiments, the prime editor fusion protein and the inhibitor of the DNA mismatch repair pathway are delivered to a cell as proteins directly. In certain embodiments, the prime editor is fused to the inhibitor of the DNA mismatch repair pathway, and the fusion protein is delivered directly into a cell.

[680] Exemplary delivery strategies include vector-based strategies, PE ribonucleoprotein complex delivery, and delivery of PE by mRNA methods. In some embodiments, the method of delivery provided comprises nucleofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycation or lipid:nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA.

[681] Exemplary methods of delivery of nucleic acids include lipofection, nucleofection, electoporation, stable genome integration (e.g., piggybac), microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycation or lipid:nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Lipofection is described in e.g., U.S. Pat. Nos. 5,049,386, 4,946,787; and 4,897,355) and lipofection reagents are sold commercially (e.g., Transfectam™, Lipofectin™ and SF Cell Line 4D-Nucleofector X Kit™ (Lonza)). Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides include those of Feigner, WO 91/17424; WO 91/16024. Delivery may be to cells (*e.g.*, *in vitro* or *ex vivo* administration) or target tissues (*e.g.*, *in vivo* administration). Delivery may be achieved through the use of RNP complexes.

[682] The preparation of lipid:nucleic acid complexes, including targeted liposomes such as immunolipid complexes, is well known to one of skill in the art (see, *e.g.*, Crystal, Science 270:404-410 (1995); Blaese et al., Cancer Gene Ther. 2:291-297 (1995); Behr et al., Bioconjugate Chem. 5:382-389 (1994); Remy et al., Bioconjugate Chem. 5:647-654 (1994); Gao et al., Gene Therapy 2:710-722 (1995); Ahmad et al., Cancer Res. 52:4817-4820 (1992); U.S. Pat. Nos. 4,186,183, 4,217,344, 4,235,871, 4,261,975, 4,485,054, 4,501,728, 4,774,085, 4,837,028, and 4,946,787).

[683] In other embodiments, the method of delivery and vector provided herein is an RNP complex. RNP delivery of fusion proteins markedly increases the DNA specificity of base editing. RNP delivery of fusion proteins leads to decoupling of on- and off-target DNA editing. RNP delivery ablates off-target editing at non-repetitive sites while maintaining on-target editing comparable to plasmid delivery, and greatly reduces off-target DNA editing even at the highly

repetitive VEGFA site 2. See Rees, H.A. et al., Improving the DNA specificity and applicability of base editing through protein engineering and protein delivery, Nat. Commun. 8, 15790 (2017), U.S. Patent No. 9,526,784, issued December 27, 2016, and U.S. Patent No. 9,737,604, issued August 22, 2017, each of which is incorporated by reference herein.

[684] Additional methods for the delivery of nucleic acids to cells are known to those skilled in the art. See, for example, US 2003/0087817, incorporated herein by reference.

[685] Other aspects of the present disclosure provide methods of delivering the prime editor constructs into a cell to form a complete and functional prime editor within a cell. For example, in some embodiments, a cell is contacted with a composition described herein (*e.g.*, compositions comprising nucleotide sequences encoding the split Cas9 or the split prime editor or AAV particles containing nucleic acid vectors comprising such nucleotide sequences). In some embodiments, the contacting results in the delivery of such nucleotide sequences into a cell, wherein the N-terminal portion of the Cas9 protein or the prime editor and the C-terminal portion of the Cas9 protein or the prime editor are expressed in the cell and are joined to form a complete Cas9 protein or a complete prime editor.

[686] It should be appreciated that any rAAV particle, nucleic acid molecule or composition provided herein may be introduced into the cell in any suitable way, either stably or transiently. In some embodiments, the disclosed proteins may be transfected into the cell. In some embodiments, the cell may be transduced or transfected with a nucleic acid molecule. For example, a cell may be transduced (*e.g.*, with a virus encoding a split protein), or transfected (*e.g.*, with a plasmid encoding a split protein) with a nucleic acid molecule that encodes a split protein, or an rAAV particle containing a viral genome encoding one or more nucleic acid molecules. Such transduction may be a stable or transient transduction. In some embodiments, cells expressing a split protein or containing a split protein may be transduced or transfected with one or more guide RNA sequences, for example in delivery of a split Cas9 (e.g., nCas9) protein. In some embodiments, a plasmid expressing a split protein may be introduced into cells through electroporation, transient (*e.g.*, lipofection) and stable genome integration (*e.g.*, piggybac) and viral transduction or other methods known to those of skill in the art.

EXAMPLES

Example 1: Enhanced Prime Editing Systems Through Identification and Manipulation of Cellular Determinants of Editing Outcomes

Introduction

[687] The ability to manipulate the genome in a programmable manner has illuminated biology and shown promise in the clinical treatment of genetic diseases. Towards the goal of expanding the breadth of sequence changes that can be precisely installed in living cells, prime editing was developed, a gene editing approach that enables all types of targeted DNA base pair substitutions, small insertions, small deletions, and combinations thereof without requiring double-strand DNA breaks or donor DNA templates (Anzalone *et al.*, 2020; Anzalone *et al.*, 2019). Prime editing has been broadly applied to introduce genetic changes in flies (Bosch *et al.*, 2021), rice and wheat (Lin *et al.*, 2020), zebrafish (Petri *et al.*, 2021), mouse embryos (Liu *et al.*, 2020), post-natal mice (Liu *et al.*, 2021), human stem cells (Sürün *et al.*, 2020), and patient-derived organoids (Schene *et al.*, 2020). Despite its versatility, prime editing efficiency can vary widely across different edit classes, target loci, and cell types (Anzalone *et al.*, 2019). To maximize the utility of prime editing, cellular determinants of prime editing outcomes were identified, and the resulting insights were used to develop prime editing systems with improved efficiency and outcome purity.

[688] The PE2 prime editing system comprises two components: an engineered reverse transcriptase (RT) fused to a Cas9 nickase (the PE2 protein) and a prime editing guide RNA (pegRNA) that contains both a spacer sequence complementary to target DNA and a 3′ extension encoding the desired edit (Anzalone *et al.*, 2020; Anzalone *et al.*, 2019) (FIG. 40A). Together, these components form a PE2–pegRNA complex that binds one strand of a target DNA locus and nicks the opposite strand. This nick exposes a DNA 3′ end that can hybridize to the primer binding site (PBS) in the pegRNA extension. Reverse transcription of the RT template within the pegRNA extension then generates a 3′ DNA flap that contains the edited sequence and ultimately leads to installation of that sequence into the genome.

[689] The PE3 system differs from PE2 by using an additional single guide RNA (sgRNA) to nick the non-edited strand away from the pegRNA target, which enhances editing efficiency by stimulating replacement of the non-edited strand (Anzalone *et al.*, 2019) (FIG. 40A).

[690] According to the current model, the newly synthesized 3′ flap displaces an adjacent strand of genomic DNA through flap interconversion (FIG. 40A). Excision of the displaced 5′ flap then allows ligation of the edited sequence into the genome. Nicking the non-edited strand in the PE3 system is thought to induce cellular replacement of the non-edited strand during heteroduplex resolution, and thus promotes copying of the edited sequence to the complementary strand.

[691] The inventors studied the roles of DNA repair mechanisms in prime editing and the development of improved prime editing systems through manipulation of those processes. As presented herein, pooled CRISPR interference (CRISPRi)-based screens were used to systematically probe the effect of 476 genes involved in DNA repair and associated processes on substitution prime editing outcomes. Specific DNA mismatch repair (MMR) genes were identified that strongly suppress prime editing efficiency and promote indel formation. Consistent with a model in which MMR reverts heteroduplex DNA formed during prime editing, classes of prime edits were identified that are less vulnerable to MMR activity and are therefore generated more efficiently, including G•C-to-C•G transversions and substitutions of three or more contiguous bases. Integrating these findings, novel prime editing systems were developed with improved editing outcomes through transient expression of a dominant negative MMR protein (MLH1dn). In MMR-proficient cell types, including induced pluripotent stem cells and primary T cells, these PE4 (PE2+MLH1dn) and PE5 (PE3+MLH1dn) systems enhance editing efficiency over PE2 and PE3 by an average of 7.7-fold and 2.0-fold, respectively, and increase edit:indel ratios (outcome purity) by 3.4-fold. Transient co-expression of MLH1dn did not result in detected changes to microsatellite repeat length, a clinically used biomarker of MMR proficiency (Umar *et al.*, 2004). By varying codon usage, nuclear localization signal, Cas9 domain mutations, and linker composition and length, an optimized prime editor architecture (PEmax – see FIG. 54B) was engineered that further increases editing efficiency in synergy with improvements from PE4, PE5, and recently developed engineered pegRNAs (epegRNAs) (Nelson *et al.*, 2021). Finally, it was shown that strategic installation of additional silent mutations nearby an intended edit can improve prime editing efficiency by weakening MMR recognition. These findings deepen the understanding of prime editing and establish prime editing systems with substantially improved efficiency and outcome purity across 169 different edits at 19 loci in seven mammalian cell types.

Results

Design of a pooled CRISPRi screen for prime editing outcomes

[692] A new genetic screening approach called Repair-seq was used to study prime editing efficiency and outcome, including the original sequence, the desired edit, and indel frequency. Repair-seq measures the effects of many loss-of-function perturbations on the outcomes of genome editing experiments by linking the identity of CRISPRi sgRNAs to edited sites in pooled screens (Hussmann *et al.*, 2021) (FIG. 40B). Briefly, a library of sgRNAs are lentivirally transduced into cells expressing the CRISPRi effector protein (dCas9–KRAB) such that most infected cells receive only one sgRNA, causing the knockdown of one gene per cell. A target sequence for genome editing is also delivered with the library sgRNAs in the same lentiviral cassette. After genome editing occurs at this target site, paired-end sequencing then enables the frequency of each editing outcome to be measured alongside each linked CRISPRi perturbation.

[693] To adapt Repair-seq for prime editing, prime editing and CRISPRi, which typically also relies on *Streptococcus pyogenes* Cas9 (SpCas9), were first made orthogonal. An SaPE2 prime editor variant was constructed by replacing the SpCas9 nickase domain in PE2 with *Staphylococcus aureus* Cas9 (SaCas9) N580A nickase (Ran *et al.*, 2015). Prime editing activity was verified using SaPE2 and orthogonal *S. aureus* pegRNAs (Sa-pegRNAs) (FIG. 47A).

[694] Next, a lentiviral Repair-seq vector was designed for screening with SaPE2 by adding a composite SaPE2 edit site. This site was comprised of a single target protospacer that was identified to be efficiently edited in HEK293T cells (FIG. 47A) and two flanking protospacers that allowed complementary-strand (the non-edited strand) nicks 50-bp downstream (+50 nick) or upstream (–50 nick) of the target (FIG. 40B, FIG. 47B). This design supported prime editing with SaPE2 in three approaches: PE2, PE3 with a +50 nick (PE3+50), or PE3 with a –50 nick (PE3–50) (FIG. 47C).

[695] To enable screening, Sa-pegRNAs were optimized for editing this site in a validated HeLa CRISPRi cell line (Gilbert *et al.*, 2013). Transfection of these cells with SaPE2, Sa-pegRNA, and Sa-sgRNA plasmids that program a G•C-to-C•G transversion at the edit site yielded up to 5.2% editing and 0.17% indels (PE2), 12% editing and 5.8% indels (PE3+50), and 2.6% editing and 19% indels (PE3–50) (FIG. 47D). The high proportion of indels from PE3–50 is consistent with past reports that 3′ overhangs left by dual nicks lead to a higher frequency of

insertions, as compared to 5′ overhangs (Bothmer *et al.*, 2017). Intended editing frequencies were increased to 9.4% with PE2, 15% with PE3+50, and 3.5% with PE3–50 using a SaPE2 construct carrying a blasticidin selection marker to enrich successfully transfected cells (FIG. 47E). Together, these efforts established a screening assay with a baseline of prime editing outcomes in an efficiency range well-suited to detect increases or decreases in editing from CRISPRi perturbations.

Identification of DNA repair components that affect prime editing outcomes

[696] With this optimized assay, Repair-seq screens of prime editing outcomes were performed with PE2 and PE3+50 in K562 and HeLa cells and with PE3–50 in HeLa cells. A library of 1,513 sgRNAs targeting 476 genes enriched for roles in DNA repair and associated processes (FIG. 48F), along with 60 non-targeting control sgRNAs, were cloned into the lentiviral screening vector (Hussmann et al., 2021) (Table 5). This library was transduced into human K562 and HeLa CRISPRi cell lines (Gilbert *et al.*, 2014; Gilbert *et al.*, 2013) and, after 5 days, cells were transfected with SaPE2, Sa-pegRNA, and Sa-sgRNA plasmids that program a G•C-to-C•G transversion at the pre-validated edit site. Genomic DNA was extracted from cells 3 days after transfection, a 453-bp region containing the CRISPRi sgRNA, edit site, and complementary nick sites was amplified by PCR, and paired-end sequencing was performed to measure the distribution of editing outcomes for each genetic perturbation (FIG. 40B, FIG. 47B). To interpret the resulting data, the frequencies of editing outcomes from cells containing a gene-targeting CRISPRi sgRNA were compared to the corresponding frequencies from cells containing non-targeting sgRNA controls. Reduction in an outcome's frequency upon gene knockdown suggests that the gene's activity promotes formation of the outcome, while an increase in frequency suggests that the gene's activity suppresses the outcome.

[697] The effect of gene knockdowns on the frequency of the intended G•C-to-C•G edit was examined. In cells with non-targeting CRISPRi sgRNAs, 4.3–4.9% (K562) and 8.5–8.7% (HeLa) of sequencing reads contained exactly the intended edit following PE2 editing (FIGs. 40C–40D). These levels increased to 14–16% (K562) and 14–16% (HeLa) for PE3+50, but decreased to 2.1–2.2% (HeLa) for PE3–50. Across all prime editing conditions screened, CRISPRi targeting of *MSH2*, *MSH6*, *MLH1*, and *PMS2*, components of the MutSα–MutLα MMR complex (Iyer *et al.*, 2006; Kunkel and Erie, 2005; Li, 2008), substantially increased editing efficiency by up to

5.8-fold for PE2, 2.5-fold for PE3+50, and 2.0-fold for PE3–50 (FIGs. 40E–40G and 47G-47J, Table 6). Knockdown of *EXO1*, an exonuclease with a role in MMR (Genschel *et al.*, 2002), also increased intended editing efficiency by up to 2.3-fold for PE2 in K562 cells. In contrast, knockdown of *LIG1*, a nick-sealing DNA ligase (Pascal *et al.*, 2004), and of *FEN1*, a 5′ flap endonuclease (Liu *et al.*, 2004), reduced the frequency of intended editing, consistent with their previously proposed roles in nick ligation and 5′ flap excision during prime editing (Anzalone *et al.*, 2019). Together, these data suggest that MMR activity antagonizes the installation of point mutations by prime editing.

[698] In addition to the intended edit, Repair-seq screens also measure the effects of genetic perturbations on the formation of editing byproducts. These byproducts were classified into four primary categories: deletions (FIG. 41A), tandem duplications (FIG. 41B), and two classes of outcomes containing unintended sequence from the pegRNA (FIGs. 41C-41D). Low baseline frequencies of total unintended edits from PE2 (0.31% in K562, 0.60% in HeLa; FIG. 41E), but more frequent and diverse unintended byproducts from PE3–50 (58% in HeLa; FIG. 48A) and PE3+50 (8.2% in K562, 9.5% in HeLa; FIG. 41F) were observed. The baseline frequencies and genetic modulators of these categories varied across PE2, PE3+50, and PE3–50 screens, providing a rich set of observations of how different prime editing configurations are processed (FIGs. 48A–48I). Two of these observations informed models for the role of MMR activity during prime editing.

[699] First, one unintended outcome category contained the intended G•C to C•G edit as well as an additional base substitution and a 1-nucleotide (nt) insertion near the target site (FIG. 41C). The sequence at these additional mutations perfectly matched 9-nt at the 3′ end of the pegRNA scaffold sequence, consistent with reverse transcription into the pegRNA scaffold and incorporation of the resulting 3′ DNA flap into partially homologous genomic sequence. Recoding the pegRNA scaffold to avoid sequence homology with the genomic target strongly reduced the frequency of these additional mutations (FIGs. 49A–49B), suggesting a generalizable approach to eliminate this type of editing byproduct. Notably, it was observed that knockdown of MMR genes increased the frequency of this editing byproduct from 0.08% to 2.0% in PE2 K562 screens (FIG. 41E, FIG. 48D). MMR thus suppresses the formation of this outcome to a larger extent than the intended edit, demonstrating that distinct prime editing intermediates can differ in the extent to which they are processed by MMR.

[700] Second, MMR knockdown reduced the frequency of most categories of unintended outcomes from PE3+50 (FIGs. 41F-41H and 48H), suggesting that transiently inhibiting some MMR activities may increase both the efficiency and outcome purity of prime editing. While tandem duplications of sequence between the nicks were common for PE3–50 (FIG. 48A), these duplication outcomes were rarer for PE3+50 (0.37% of non-targeting reads in K562, 2.3% in HeLa) and were reduced by up to 3.7-fold (K562) and 1.5-fold (HeLa) by MMR knockdown (FIG. 41F, FIG. 48H). The most abundant class of unintended PE3+50 outcomes contained sequence from the reverse-transcribed 3′ DNA flap that does not rejoin genomic sequence at the intended flap annealing location (5.1% of non-targeting reads in K562, 3.8% in HeLa; FIG. 41D). In both cell types, the frequency of unintended flap rejoining outcomes increased upon knockdown of *HLTF*, a fork remodeling helicase (Poole and Cortez, 2017), but was reduced by up to 1.7-fold by knockdown of MMR genes (FIG. 41H). Finally, MMR knockdown substantially reduced deletions from PE3+50 by up to 3.7-fold (K562) and 1.6-fold (HeLa; FIG. 41G). MMR knockdown in K562 cells also qualitatively shifted the observed boundaries of deleted sequence. For both PE3 configurations, genomic sequence between the two SaPE2-induced nicks was most frequently deleted, but deletions extending outside of this region were also observed (FIG. 41I, top). MMR knockdown decreased the frequency of these longer deletions dramatically more than deletions between the programmed nicks with PE3+50 (FIG. 41I, bottom and FIG. 41J), suggesting that MMR activity may cause the formation of longer deletions during prime editing.

Model for mismatch repair of prime editing intermediates

[701] The effects of MMR knockdown on both intended and unintended editing outcomes in these Repair-seq screens led to a working model for the role of MMR during prime editing. In eukaryotes, MMR resolves DNA heteroduplexes containing a single base mismatch or small insertion-deletion loop (IDL) by selectively replacing the DNA strand that contains a nearby nick (Iyer *et al.*, 2006; Kunkel and Erie, 2005; Li, 2008). To initiate MMR, the heteroduplex is first bound by MutSα (MSH2–MSH6), which recognizes base mismatches and 1- to 2-nt IDLs (Warren *et al.*, 2007), or by MutSβ (MSH2–MSH3), which recognizes 2- to 13-nt IDLs (Gupta *et al.*, 2012) (FIG. 42C). Next, MSH2 recruits the MutLα heterodimer (PMS2–MLH1), which incises only the nick-containing strand around the heteroduplex (Fang and Modrich, 1993;

Kadyrov *et al.*, 2006; Pluciennik *et al.*, 2010; Thomas *et al.*, 1991). From these incisions, EXO1 mediates 5′-to-3′ excision of the heteroduplex (Genschel *et al.*, 2002), polymerase δ resynthesizes the excised DNA strand, and ligase I (LIG1) seals the nascent strand to complete repair (Iyer *et al.*, 2006; Kunkel and Erie, 2005; Zhang *et al.*, 2005).

[702] Without wishing to be bound by theory, MMR may engage a specific prime editing intermediate: a DNA heteroduplex formed by hybridization of the reverse-transcribed 3′ DNA flap to the adjacent genomic DNA (FIG. 42A). If MutSα–MutLα recognizes the heteroduplex within this structure, the 3′ nick present after flap equilibration—but before ligation—could stimulate selective excision of the edited strand and subsequent repair to regenerate the original, unedited sequence. Alternatively, MMR may also prevent productive flap interconversion by rejecting annealing of the edited 3′ flap to the genomic target (Sugawara *et al.*, 2004). In either case, inhibiting MMR during prime editing could delay heteroduplex repair or increase the likelihood of nick ligation. Moreover, successful nick ligation would remove the ability of MMR to bias resolution of the heteroduplex towards removal of the edited product. This model is supported by substantial increases in PE2 editing efficiencies (1.6- to 5.8-fold) from knockdown of MutSα–MutLα genes (FIG. 40C–40D), which strongly suggest that the 3′-nicked heteroduplex intermediate is often repaired by MMR to the original sequence before ligation of the nick (FIG. 42A). Interfering with MMR reversion of these intermediates can thus enhance prime editing efficiency.

[703] This model also explains the benefit from complementary-strand nicks in the PE3 editing system. Nicking the non-edited strand of the heteroduplex intermediate introduces an additional strand discrimination signal that could direct MMR to more frequently replace the non-edited strand, leading to higher prime editing efficiency and dampened effects of MMR suppression (FIG. 42B). Consistent with this model, knockdown of MutSα–MutLα components increased intended G•C-to-C•G editing by up to 5.8-fold for PE2, but only by up to 2.6-fold for PE3+50 (FIG. 40G). MMR activity may also promote successful prime editing by directing excision of the non-edited strand, but only for a prime editing intermediate in which the edited strand has been ligated and the non-edited strand is nicked. However, MMR knockdown substantially enhances PE3+50 editing in K562 and HeLa cells (FIG. 40F), suggesting that this intermediate is uncommon and that MMR typically repairs the heteroduplex before nick ligation.

[704] The mechanism of MMR may also explain how MutSα–MutLα gene knockdown reduces indel byproducts from PE3+50 (FIG. 41F). During repair of prime editing heteroduplex intermediates, MutLα may induce DSBs by indiscriminately nicking the target locus, particularly when both DNA strands already contain pegRNA- and sgRNA-programmed nicks (FIG. 49D). DSBs formed after excision from these additional, non-programmed nicks could broaden the boundaries of indels at prime edited loci. In agreement with this hypothesis, knockdown of MutSα–MutLα components disproportionately reduced PE3+50 deletion outcomes outside of the sequence between pegRNA and sgRNA nicks (FIG. 41I). Taken together, these findings from Repair-seq screens thus support a model in which MMR activity strongly suppresses intended prime editing outcomes and instead promotes indel byproducts.

MMR inhibition improves prime editing at endogenous loci

[705] The effect of MMR on prime editing was tested with canonical SpCas9-based prime editors at endogenous genomic loci and in additional cell types. HEK293T cells were treated with siRNAs targeting *MSH2*, *MSH6*, *MLH1*, or *PMS2*, the cells were cultured for 3 days to allow siRNA-mediated knockdown (FIG. 49E), then plasmids encoding PE2 and pegRNAs that program point mutations at three endogenous genomic loci were transfected (*EMX1*, *RUNX1*, and HEK293T cell site 3, hereafter referred to as *HEK3*). Across these loci, it was observed that mRNA knockdown strongly increased average PE2 editing from 7.7% to 25% with a decrease in indel frequency from 0.39% to 0.28% (FIG. 42C), but improved average PE3 editing efficiency to a lesser extent (from 25% to 37%). These results are consistent with the model presented herein, which states that complementary-strand nicking improves prime editing efficiency by directing MMR of the unedited strand. Thus, the impact of MMR on PE3 editing efficiency is tempered by its opposing effects on reverting the 3′ flap intermediate (which impedes prime editing) and on mediating replacement of the unedited strand (which promotes prime editing). Additionally, knockdown of MMR genes reduced the frequency of PE3 indels from 5.5% down to 3.2% on average, resulting in a 2.9-fold increase in PE3 outcome purity (FIG. 42C). These findings demonstrate that inhibition of MMR components can improve prime editing outcomes at endogenous genomic sites in human cells.

[706] Prime editing was measured in MMR-deficient Δ*MSH2* or Δ*MLH1* haploid HAP1 cells. PE2 prime editing efficiency was much greater in MMR-deficient HAP1 cells (17% at *HEK3*,

5.0% at *EMX1*) than in wild-type control cells (0.44% at *HEK3*, 0.07% at *EMX1*; FIG. 42D). Moreover, consistent with the hypothesis stated herein, nicking the unedited strand at these loci did not affect editing efficiency in MMR-deficient HAP1 cells (FIG. 42D). Taken together, these results further support a model in which MMR impedes prime editing by promoting excision of the edited DNA strand, even though this effect is partially counterbalanced in the PE3 system by the role of MMR in replacing the non-edited strand.

Engineered dominant negative MMR proteins enhance prime editing efficiency and precision

[707] Encouraged that cellular pre-treatment with MMR-targeting siRNAs can enhance prime editing efficiency, strategies for simultaneous co-delivery of prime editors and MMR-inhibiting agents were explored. Co-transfection of PE2 and *MLH1* siRNAs without pre-treatment did not substantially increase editing efficiency after 3 days (FIG. 49F). It was hypothesized that dominant negative MMR protein variants could instead be transiently co-expressed with PE2 or as fusion proteins with PE2 to enhance prime editing. HEK293T cells were co-transfected with plasmids encoding PE2, pegRNAs, and ATPase-deficient mutants of human MSH2, MSH6, PMS2, and MLH1 (Iaccarino *et al.*, 1998; Räschle *et al.*, 2002; Tomer *et al.*, 2002), or endonuclease-deficient mutants of PMS2 and MLH1 (Gueneau *et al.*, 2013; Kadyrov *et al.*, 2006) (FIG. 43A). Of these mutants, ATPase-impaired MLH1 E34A and endonuclease-impaired MLH1 Δ756 increased PE2 editing efficiency by 1.6- to 3.1-fold for three single-base substitution edits at the *HEK3*, *EMX1*, and *RUNX1* loci.

[708] Next, additional dominant negative MLH1 variants were engineered and tested to maximize the enhancement of prime editing efficiency. The MLH1 N-terminal domain (NTD) mediates MutLα (PMS2–MLH1) recruitment to MSH2 during MMR (Plotz *et al.*, 2003) and contains an ATPase essential for MutLα function (Kadyrov *et al.*, 2006) (FIG. 43B). In contrast, the MLH1 C-terminal domain (CTD) dimerizes with PMS2 and contributes to MutLα endonuclease activity critical for MMR (Gueneau *et al.*, 2013) (FIG. 43B). While the MLH1 Δ756 dominant negative variant disrupts this endonuclease, it was found that a larger deletion of these residues (MLH1 Δ754-756) further elevated prime editing efficiency at three sites tested (FIG. 43C, FIG. 50A). Combining ATPase and endonuclease mutations (MLH1 E34A Δ754-756) did not further improve prime editing, however (FIG. 43C). Comparing these dominant negative MLH1 variants across ten prime edits, it was observed that MLH1 Δ754-756 enhanced

PE2 editing efficiencies to the greatest extent on average (by 3.2-fold; FIG. 50D). Thus, MLH1 Δ754-756 was designated as MLH1dn. MLH1dn improved prime editing in a dose-dependent manner within HEK293T cells (FIG. 50B). MLH1dn did not increase editing in MMR-deficient HCT116 cells (Parsons et al., 1993) (FIG. 50C). Both human and mouse MLH1dn also improved substitution prime editing efficiency across three sites in human HEK293T cells and four sites in mouse N2A cells (FIGs. 50D– 50E).

[709] Next, shorter MLH1 truncations that can also inhibit MMR during prime editing were identified. On its own, the MLH1 NTD (residues 1-335) modestly enhanced PE2 editing by 1.5- to 1.8-fold, but introducing the E34A ATPase mutation weakened this improvement at two of three tested loci (FIG. 43C, FIG. 50A). Appending a nuclear localization signal (NLS) to the MLH1 NTD boosted PE2 editing by 1.9- to 2.5-fold, to a similar degree as full-length MLH1dn. In contrast, the MLH1 CTD (residues 501-756), as well as endonuclease mutants and NLS-appended variants thereof, did not substantially enhance PE2 editing (FIG. 43C, FIG. 50A). These data suggest that dominant negative MLH1 variants can inhibit MMR by forming catalytically impaired MutLα complexes with PMS2 or by saturating the binding of MSH2, which would prevent productive recruitment of MutLα during MMR. Full-length MLH1dn, which can do both, could inhibit MMR through both mechanisms, but NLS-appended MLH1 NTD (hereafter referred to as $MLH1^{NTD}$–NLS) would be expected to inhibit MMR through MSH2 binding as it lacks the CTD needed to dimerize with PMS2 (Guerrette *et al.*, 1999). $MLH1^{NTD}$–NLS improves PE2 editing efficiency to almost the same extent as MLH1dn (FIG. 43C, FIG. 50A), suggesting that MSH2 sequestration is an effective mechanism for inhibiting MMR.

[710] MLH1dn or $MLH1^{NTD}$–NLS were also directly fused to the PE2 protein (via a 32aa (SGGS)x2-XTEN16-(SGGS)x2 linker), but did not exhibit higher editing efficiency compared to PE2 alone (FIG. 50A). However, appending MLH1dn to PE2 with a self-cleaving P2A linker (PE2–P2A–MLH1dn) (Kim *et al.*, 2011) enhanced prime editing efficiency at three sites by 2.0- to 2.7-fold (FIG. 50A). Among the 55 total dominant negative MMR protein variants tested across three loci in HEK293T cells, PE2 and MLH1dn expressed *in trans* provided the greatest average enhancement in PE2 editing efficiency (3.2-fold). Strong average improvement of PE2 editing was also observed from PE2 with $MLH1^{NTD}$–NLS expressed *in trans* (2.7-fold) and PE2–P2A–MLH1dn (2.4-fold) (FIG. 43C). These three variants also increased PE3 editing efficiency

by 1.2-fold on average but reduced undesired indel products by 1.4- to 4.0-fold (FIG. 43E). PE2 editing with MLH1dn was designated as the PE4 system, and PE3 editing with MLH1dn as the PE5 system (FIG. 43F). $MLH1^{NTD}$–NLS can also offer enhanced prime editing efficiencies with a smaller protein (355-aa) compared to MLH1dn (753-aa). When a single construct is required, the use of PE2–P2A–MLH1dn is recommended.

[711] The generality of PE4 and PE5 systems were assessed with seven additional single-base substitution edits across different genomic loci in HEK293T cells. On average, PE4 improved editing efficiency over PE2 by 2.0-fold with minimal indels (<0.4% on average; FIG. 43G). PE5 improved editing over PE3 by 1.2-fold and enhanced edit:indel purity by 3.0-fold (FIG. 43G). Whether MLH1dn could elevate the efficiency of PE3b, a prime editing strategy that uses a complementary-strand nick specific for the edited sequence to minimize coincident nicks on both strands that promote indel formation (Anzalone *et al.*, 2019), was also tested. PE3b with co-expression of MLH1dn (hereafter referred to as PE5b) increased editing at the *FANCF* locus by 1.7-fold over PE3b with low indel frequencies (<0.6%; FIG. 50H). Moreover, it was shown that PE4 offers substantially improved prime editing performance compared to PE2 when indels must be minimized or at loci where complementary-strand nicks yield unproductive editing outcomes (FIG. 50H). Prime editing enhancement from MLH1dn versus complete *MLH1* knockout was also compared in clonal HeLa cell lines. At four sites tested, *MLH1* knockout enhanced PE2 and PE3 editing to a larger degree than MLH1dn co-expression (PE4 and PE5; FIG. 50F), suggesting opportunities for additional prime editing enhancement through further modulation of this pathway. Collectively, these data establish PE4 and PE5 systems that substantially enhance prime editing efficiency and outcome purity at a variety of endogenous genomic loci in human cells.

Characterization of the types of prime edits enhanced by PE4 and PE5

[712] Next, the extent to which MLH1dn improves prime editing across a wide range of different edit types was studied. PE4 was compared with PE2 in HEK293T cells using 84 pegRNAs that together introduce all 12 possible single-base substitutions across seven genomic loci. Among these edits, MLH1dn improved prime editing efficiency by 2.0-fold and reduced average indel frequencies from 0.40% to 0.31% compared to PE2 (FIG. 44A–B, FIGs. 51A–51B).

[713] MLH1dn increased the efficiency of A•T-to-G•C, T•A-to-G•C, and C•G-to-A•T prime edits to a lesser extent, by 1.7-fold over PE2. Notably, G•C-to-C•G substitutions, which form C•C mismatches after 3′ flap hybridization, were the least improved with MLH1dn (1.2-fold), consistent with previous studies establishing that C•C mismatches are not efficiently repaired by MMR (Lahue *et al.*, 1989; Su *et al.*, 1988; Thomas *et al.*, 1991). This finding suggests that G•C-to-C•G edits more effectively evade MMR and may therefore yield higher basal editing efficiency. Consistent with this possibility, G•C-to-C•G edits with PE2 were substantially more efficient (27%) than G•C-to-A•T (18%) or G•C-to-T•A (20%) edits among PAM-altering prime edits across seven endogenous loci (FIG. 51C).

[714] To further confirm that G•C-to-C•G substitutions are less sensitive to MMR, the effect of MMR knockdown or knockout was tested on the efficiency of a G•C-to-C•G prime edit at the *RNF2* locus that is unaffected by MLH1dn co-expression (FIG. 51A). siRNA knockdown of MMR components in HEK293T cells (FIG. 51D) and knockout of *MSH2* or *MLH1* in HAP1 cells (FIG. 51E) did not change the efficiency of this G•C-to-C•G edit. G•C-to-A•T, G•C-to-C•G, and G•C-to-T•A edits were also compared with SaPE2 at the pre-validated screening site in HeLa CRISPRi cells (FIG. 40B). PE2 and PE3+50 more efficiently installed the G•C-to-C•G edit than the G•C-to-A•T or G•C-to-T•A edits, consistent with weaker MMR activity at C•C mismatches (FIG. 51F). Furthermore, CRISPRi knockdown of *MSH2* improved G•C-to-A•T and G•C-to-T•A editing efficiencies (by 16-fold for PE2 and 4.3-fold for PE3+50) to a greater extent than G•C-to-C•G editing (by 4.0-fold for PE2 and 1.9-fold for PE3+50). These data strongly support that G•C-to-C•G prime edits are less susceptible to repair by MMR and are consequently installed with higher efficiency.

[715] PE5 and PE3 were also compared with the same set of 84 pegRNAs used in the above experiments. PE5 yielded an average 1.2-fold increase in editing relative to PE3 (FIG. 44A, FIG. 51A, FIG. 51G). MLH1dn substantially reduced the frequency of unwanted indel products by 2.2-fold on average, resulting in 2.8-fold higher edit:indel purity (FIG. 44A). Collectively, this analysis of 84 different single-nucleotide substitutions at seven genomic loci strongly support a model in which PE4 and PE5 systems improve prime editing efficiency and outcome purity by impairing counterproductive MMR of prime editing intermediates.

[716] To determine if MLH1dn could also improve small insertion and deletion prime edits, 1- and 3-bp insertions and deletions were next installed with PE4 and PE5 in HEK293T cells.

Across 12 pegRNAs at three loci, it was observed that PE4 increased average editing efficiency by 2.2-fold over PE2 with no increase in indel frequency, while PE5 increased editing efficiency by 1.2-fold and edit:indel purity by 2.9-fold over PE3 (FIG. 44C and 51I). To evaluate the impact of MLH1dn on larger sequence changes, PE2 and PE4 were tested with a combined 33 pegRNAs that together program 1-, 3-, 6-, 10-, 15-, and 20-bp insertions and deletions at the *HEK3* and *FANCF* loci. It was found that the enhancement in prime editing efficiency from MLH1dn gradually declined as the length of the insertions and deletions increased (FIGs. 44D and 51H), such that 15- and 20-bp insertions and deletions were edited by PE4 only 1.1-fold more efficiently on average than by PE2. These observations are consistent with previous reports that MMR repairs IDLs ≤13-nt in length (Acharya *et al.*, 1996; Genschel *et al.*, 1998; Umar *et al.*, 1994). These results together demonstrate that PE4 and PE5 strategies can enhance small (<15-bp) targeted insertions and deletions.

Installing additional silent mutations can increase prime editing efficiency by evading MMR

[717] Whether other classes of prime edits could bypass MMR was also explored. In addition to point mutations, short insertions, and short deletions, prime editing can also install multiple neighboring or contiguous base substitutions (Anzalone *et al.*, 2019). MutSα and MutSβ heterodimers each recognize specific heteroduplex DNA structures (Gupta *et al.*, 2012; Warren *et al.*, 2007), suggesting that a DNA bubble of contiguous mismatches could weaken recognition from these MMR components. To assess this possibility, PE2 and PE4 were tested with 35 different edits that generate 1–5 contiguous base substitutions at five genomic loci in HEK293T cells. Across seven edits that mutate two adjacent bases, PE4 yielded 2.3-fold higher editing efficiency than PE2, comparable to the 2.4-fold enhancement for single-base substitutions at the same target nucleotides (FIG. 44E, FIG. 52A). In contrast, the editing frequencies of longer 3- to 5-base contiguous substitutions were only improved by 1.2- to 1.5-fold with PE4 relative to PE2. The reduced impact of MMR on these larger edits was reflected in higher average PE2 editing efficiency of for 3- to 5-bp contiguous substitutions (13% across 21 edits) compared to 1–2 bp contiguous substitutions (4.8% across 14 edits) (FIG. 44F, FIG. 52A). Inhibiting MMR with MLH1dn (PE4) raised the average editing frequency of these 3–5-bp and 1–2-bp contiguous substitutions to 16% and 10%, respectively, lessening the difference between these edit types.

[718] Installing additional silent mutations nearby the intended edits could similarly increase prime editing efficiency by weakening repair of the resulting heteroduplex (FIG. 51G), even in the absence of any MLH1 inhibition. PE2 was tested at six gene targets with either pegRNAs that program only a coding mutation or with pegRNAs that program additional silent mutations close to the coding edit (most fewer than 5 bp away). At four of six sites, making additional silent mutations increased the efficiency of the desired coding changes, by an average of 1.8-fold for the best pegRNAs at each site (FIGs. 44H and 52B). Inhibiting MMR with MLH1dn (PE4) improved editing efficiency with the best silent pegRNAs to a lesser extent (1.2-fold on average) than with pegRNAs that only make a coding mutation (1.7-fold), suggesting that these additional silent mutations enhance editing by evading MMR. Consistent with this mechanism, at the two sites in which silent pegRNAs do not improve editing, it was observed that the tested silent mutations do not substantially dampen the effect of MMR inhibition on editing efficiency (FIG. 52B). Collectively, these findings support that MMR less efficiently repairs heteroduplexes containing three or more contiguous mismatched bases and reveal, surprisingly, that the strategic installation of additional silent or benign mutations nearby the desired edit can increase prime editing efficiency by evading MMR reversal of prime editing intermediates, even when using PE2- or PE3-based systems that do not manipulate MMR activity.

PE4 and PE5 strongly improves prime editing in cell types with fully active MMR

[719] HEK293T cells are partially MMR-deficient due to MLH1 promoter hypermethylation (Trojan *et al.*, 2002), which in light of the above findings may explain higher prime editing efficiency observed in HEK293T cells compared to other mammalian cell types (Anzalone *et al.*, 2019). To evaluate whether MLH1dn also improves prime editing to a greater degree in cells with fully active MMR, prime editing was compared in HEK293T cells and in MMR-proficient HeLa cells using 30 pegRNAs that together target seven loci. Consistent with higher MMR activity in HeLa cells (Holmes et al., 1990; Thomas et al., 1991), complementary-strand nicking (PE3 vs. PE2) resulted in much larger improvements in prime editing efficiency in HeLa cells (22-fold on average) than in HEK293T cells (3.5-fold; FIGs. 44I-44J and 52C,). Likewise, PE4 enhanced average editing efficiency over PE2 to a much greater extent in MMR-proficient HeLa cells (6.7-fold) than in partially MMR-deficient HEK293T cells (2.0-fold) across the same 30 edits, while maintaining minimal indel frequencies (0.67% on average in HeLa cells). Similarly,

PE5 improved average editing efficiency over PE3 by 1.9-fold in HeLa cells but only by 1.1-fold in HEK293T cells. MLH1dn also increased edit:indel ratios on average by 2.8-fold in HeLa cells and 3.2-fold in HEK293T cells (FIG. 44I).

[720] PE4 and PE5 were also assessed in MMR-proficient K562 and U2OS cell lines (Matheson and Hall, 2003; Peng et al., 2014). Across six sites in K562 cells and four sites in U2OS cells, PE4 and PE5 elevated average prime editing efficiency by 6.0-fold and 1.7-fold over PE2 and PE3, respectively (FIGs. 44I–44J), again offering a much larger benefit than was observed in MMR-impaired HEK293T cells. PE5 in these cells also improved edit:indel purity over PE3 by 2.6-fold. Intriguingly, although PE4 only increased G•C-to-C•G editing frequency at *DNMT1* by 1.4-fold over PE2 in HEK293T cells (FIG. 51A), a larger improvement was observed in HeLa, K562, and U2OS cells (averaging 2.7-fold; FIG. 44J), suggesting that PE4 and PE5 can enhance prime editing in MMR-proficient cell types even for classes of edits that evade MMR activity more effectively. Together, this comparison between 70 edits across seven endogenous sites in HEK293T, HeLa, K562, and U2OS cells illustrate that MLH1dn substantially improves prime editing efficiency, but especially in cell types that are not MMR-deficient.

Effect of MLH1dn on prime editing outcome purity

[721] How MLH1dn reduces unintended prime editing outcomes at endogenous genomic targets was characterized next in depth. To decouple steps that lead to prime editing from those that lead to indel byproducts, prime editors were tested with non-editing pegRNAs that template a 3′ DNA flap with perfect complementarity to the target locus and would result in no sequence change at the target locus (FIG. 45A). Across four sites in HEK293T cells, these non-editing pegRNAs yielded an average 4.4% indels with PE3 and 4.3% indels with PE5 (FIGs. 45B–45C, FIG. 53A). In contrast, four pegRNAs that program single-nucleotide mutations at these sites generated on average 8.5% indels with PE3 and 4.8% indels with PE5. These data show that MMR inhibition with MLH1dn does not alter indel frequencies in the absence of a heteroduplex. It was also observed that MLH1dn did not alter the frequency of indels from pegRNAs and nicking sgRNAs co-transfected with PE2–dRT, a PE2 containing an inactivated reverse transcriptase, or with SpCas9 H840A nickase (nCas9; FIG. 53A), suggesting that MMR does not affect the repair of a doubly nicked intermediate that lacks a reverse-transcribed 3′ flap. These

results together strongly suggest that MMR engagement of the prime editing heteroduplex intermediate stimulates indel products that can be mitigated with PE5.

[722] Next, evidence that MMR activity expands the size of prime editing indel byproducts was looked for at seven endogenous loci in HEK293T cells. The distribution of unintended target sequence deletions from PE3 and PE5 was compared with 84 pegRNAs encoding single-base substitutions. Consistent with findings from the Repair-seq screens (FIG. 41I), MLH1dn in the PE5 system reduced deletions outside of the pegRNA- and sgRNA-programmed nicks to a greater extent than deletions between these nicks (FIGs. 45D–45E, 53B). In comparison, MLH1dn did not affect the distribution of prime editing deletion outcomes from a non-editing pegRNA that does not create a mismatch (FIG. 53C). Taken together, this analysis supports a model in which non-programmed incisions by MutLα and subsequent excision at the target locus generates larger indel byproducts during MMR of prime editing intermediates.

[723] Lastly, it was examined how MLH1dn impacts pegRNA scaffold sequence incorporation and unintended flap rejoining, two unintended outcome categories previously identified in the Repair-seq screens (FIGs. 41C–41D). Across the 84 pegRNAs tested in HEK293T cells above, PE5 reduced the average frequency of these combined outcomes by 1.6-fold compared to PE3 (from 1.8% to 1.0% on average, FIG. 45F, FIG. 53D). Consistent with data from PE2 and PE3+50 screens (FIGs. 41E–41F), these outcomes were much rarer in the absence of a complementary-strand nick (0.27% frequency for PE2, 0.28% for PE4; FIG. 53D), suggesting that they typically form through a doubly nicked intermediate. Thus, PE5 broadly narrows the size of unintended deletions and reduces the frequency of pegRNA scaffold sequence incorporation and unintended flap rejoining compared to PE3.

Effect of MLH1dn on off-target genomic DNA changes

[724] It was next assessed whether MMR component manipulation could influence off-target editing. PE2 and PE4 were tested in HEK293T cells with eight pegRNAs targeting the *HEK3*, *EMX1*, *FANCF*, and *HEK4* loci. The resulting genomic changes were measured at the top four Cas9 off-target sites identified by CIRCLE-seq for each of the targeted loci (Tsai *et al.*, 2017). The average frequency of off-target prime editing remained very low with or without ML1dn (0.094% with PE2, 0.12% with PE4), while the average efficiency of on-target editing increased from 9.7% for PE2 to 20% for PE4; (FIG. 45G, FIG. 53E). These data are consistent with

previous reports noting the high DNA specificity of prime editing (Anzalone *et al.*, 2019; Jin *et al.*, 2021; Kim *et al.*, 2020), and suggest that MLH1dn does not substantially increase off-target prime editing.

[725] Next, it was asked whether transient inhibition of MMR with MLH1dn could induce genomic mutations independent of prime editor activity. In humans, MMR deficiency is a risk factor for colorectal cancer and most frequently manifests as mutations altering the length of repetitive microsatellite sequences within the genome (Fishel *et al.*, 1993; Leach *et al.*, 1993; Parsons *et al.*, 1993). Because errors from microsatellite replication are repaired almost exclusively by MMR (Strand *et al.*, 1993; Tran *et al.*, 1997), microsatellite instability is used clinically as a measure of MMR activity (Bacher *et al.*, 2004; Umar *et al.*, 2004).

[726] Microsatellite instability was evaluated by high-throughput sequencing of 17 mononucleotide tracts previously shown to be responsive to MMR inhibition and used as clinical biomarkers for MMR deficiency (Hempelmann *et al.*, 2015). These microsatellite loci were analyzed in HAP1 cells, HeLa cells and in MMR-deficient HCT116 cells. HCT116 cells exhibited substantially shorter microsatellite lengths on average (13.9-nt) than HAP1 or HeLa cells (18.4-nt; FIG. 45H, FIG. 53F). To gauge whether this assay can also detect a shorter duration of MMR deficiency in the same cell type, microsatellite instability was compared in wild-type HAP1 cells and monoclonal HAP1 cells grown for 2 months (~60 cell divisions) following knockout of *MSH2*. These MMR knockout cells exhibited a 0.24-nt average decrease in microsatellite length (FIG. 45H, FIG. 53F), demonstrating that even recent MMR impairment can be detected through the accumulation of microsatellite length erosion.

[727] To assess the effect of transient MLH1dn expression as used in PE4 and PE5 systems, microsatellite instability was next measured in MMR-proficient HeLa cells 3 days after transfection with plasmids encoding PE2 or PE4. Although MLH1dn improved prime editing efficiency from 1.3% (PE2) to 7.6% (PE4) at the on-target locus (FIG. 51G), average microsatellite lengths were indistinguishable in PE2-treated cells compared to in PE4-treated cells (<0.01 nt of difference; FIG. 45H, FIG. 53F). These data indicate that transient MLH1dn expression can enhance prime editing without stimulating detectable instability at 17 microsatellites used for clinical diagnosis of MMR deficiency. As microsatellite repeats are up to 10^5-fold more susceptible to MMR inhibition than the rest of the genome on average (Strand *et al.*, 1993; Tran *et al.*, 1997), these findings suggest that the PE4 and PE5 editing systems that use

transient MLH1dn expression, including MLH1dn mRNA co-delivery described below, can enhance on-target editing without substantial off-target mutational burden.

PEmax systems with optimized architectures and synergy with engineered pegRNAs

[728] To further improve prime editing, the PE2 protein was optimized by varying RT codon usage, mutations within the SpCas9 domain, the length and composition of peptide linkers between nCas9 and RT, and the location, composition, and number of NLS sequences (FIGs. 56A–56B). Among 21 such variants tested, the greatest enhancement in editing efficiency was observed from a prime editor architecture that uses a human codon-optimized RT, a 34-aa linker containing a bipartite SV40 NLS (Wu *et al.*, 2009), an additional C-terminal c-Myc NLS (Dang and Lee, 1988), and R221K N394K mutations in SpCas9 previously shown to improve Cas9 nuclease activity (Spencer and Zhang, 2017) (FIG. 46A, FIG. 56A). This optimized prime editor architecture was designated as PEmax (SEQ ID NO: 99 disclosed herein; also see schematic of FIG. 54A). At seven target sites tested in HeLa cells, this optimized prime editor architecture (hereafter referred to as PEmax) outperforms other improved prime editor variants, including PE2*, which includes additional NLS sequences (Liu et al., 2021), and CMP–PE–V1, which contains high-mobility peptides (Park et al., 2021) (FIGs. 56B–56D). Inserting high-mobility peptides into PEmax (CMP–PEmax) did not further improve prime editing (FIGs. 56C–56D).

[729] Across seven substitution edits targeting different loci, using the PEmax architecture with the PE2 system (PE2max) increased the average frequency of intended editing by 2.3-fold in HeLa cells and 1.2-fold in HEK293T cells over the original PE2 architecture (FIG. 52B). Similarly, PE3 using the PEmax architecture (PE3max) increased average editing efficiency over PE3 by 2.8-fold in HeLa cells and 1.2-fold in HEK293T cells (FIGs. 53B and 56C). PE3max also slightly reduced average edit:indel purity by 1.2-fold in both cell types, which may reflect enhanced nickase activity from the Cas9 R221K and N394K mutations within the PEmax architecture (Spencer and Zhang, 2017).

[730] It was also observed that PE4max (PE4 using the PEmax architecture) enhanced average editing frequencies over PE4 by 1.9-fold in HeLa cells and by 1.1-fold in HEK293T cells (FIG. 56E). Finally, PE5max (PE5 using the PEmax architecture) similarly increased average editing efficiency over PE5 by 2.2-fold in HeLa cells and 1.2-fold in HEK293T cells.

[731] In a separate improvement to prime editing, engineered pegRNAs (epegRNAs) were recently developed, which contain an additional3′ RNA structural motif that increases prime editing efficacy (Nelson *et al.*, 2021) (FIG. 46C). To assess whether optimized PE4max and PE5max systems can synergize with epegRNAs, PE4max and PE5max were tested in combination with epegRNAs in HeLa and HEK293T cells. Across seven substitution edits targeting different loci, epegRNAs improved PE4max editing efficiency over normal pegRNAs by an average of 2.5-fold in HeLa cells and 1.5-fold in HEK293T cells (FIG. 46B). Similarly, epegRNAs enhanced PE5max editing over normal pegRNAs by 1.4-fold in HeLa cells and 1.1-fold in HEK293T cells, without affecting edit:indel purity.

[732] Combining all enhancements to prime editing systems described above—MLH1dn, PEmax, and epegRNAs—dramatically improved prime editing performance. PE4max with epegRNAs enhanced editing efficiency by an average of 72-fold in MMR-proficient HeLa cells and 3.5-fold in MMR-deficient HEK293T cells relative to PE2 with normal pegRNAs (FIG. 46B). PE5max with epegRNAs also improved editing efficiency over PE3 with pegRNAs on average by 12-fold in HeLa cells and 1.6-fold in HEK293T cells, with outcome purity increasing on average by 4.6-fold in HeLa cells and 3.3-fold in HEK293T cells. Collectively, these results demonstrate that combining PE4/PE5, PEmax, and epegRNA approaches can greatly enhance prime editing efficiency and outcome purity.

Prime editing of disease-relevant loci and cell types with PE4 and PE5

[733] To establish the applicability of these improved editing systems, PE4max and PE5max were applied to edit therapeutically relevant loci in wild-type HeLa and HEK293T cells. First, a silent G•C-to-A•T transversion was made at the 6th codon of *HBB*, which is mutated in sickle cell disease patients (Ingram, 1956). Second, the G127V allele (a G•C-to-T•A transversion) was installed in *PRNP* that confers resistance to prion disease (Asante *et al.*, 2015; Mead *et al.*, 2009). Third, a silent C•G-to-T•A mutation was introduced at a *CDKL5* site known to contain a causative mutation for CDKL5 deficiency disorder, a severe neurodevelopmental condition (Olson *et al.*, 2019). Fourth, the *CXCR4* P191A allele (a G•C-to-C•G edit) that inhibits HIV infection in human cells was installed (Liu *et al.*, 2018). Fifth, the *IL2RB* H134D Y135F (non-adjacent T•A-to-A•T and G•C-to-C•G edits) variant that enables orthogonal IL-2 receptor responsiveness for adoptive T cell transfer therapy was generated (Sockolosky *et al.*, 2018).

Lastly, the BCL11A repressor binding site within the *HBG1* and *HBG2* fetal hemoglobin gene promoters was recoded to a GATA1 transcriptional activator motif (non-adjacent G•C-to-A•T and C•G-to-A•T edits), which in principle could induce fetal hemoglobin expression for treatment of hemoglobinopathies (Amato *et al.*, 2014). Sequences are shown in Table 4, below.

[734] Table 4. Sequences Used for Prime Editing of Disease-Relevant Loci and Cell Types

Oligonucleotides	Sequence	SEQ ID NO:
CDKL5 c.1412delA correction with silent edit epegRNA	mG*mA*mG*rGrGrArCrUrCrCrUrArGrArGrGrArCrUrGrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrArUrArUrUrGrArCrArCrArArUrUrCrCrCrCrArArUrCrCrUrCrUrArGrGrArGrUrCrArArArArArArCrArCrCrGrUrCrArGrGrGrUrCrArGrGrArGrCrCrCrCrCrCrCrCrUrGrCrArCrCrCrArGrGrArArArCrCrCrUrCrArArArGrUrCrGrGrGrGrGrGrCrArA*mC*mC*mC	237
CDKL5 +86 nick sgRNA	mG*mC*mA*rGrArArCrCrGrCrCrArCrUrCrArUrUrCrArGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrUmU*mU*mU	238
CDKL5 -3 nick sgRNA (for PE3b/PE5b)	mA*mC*mA*rCrArArUrUrCrCrCrCrArArUrCrCrUrCrUrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrUmU*mU*mU	239
PRNP G127V install pegRNA	mG*mC*mA*rGrUrGrGrUrGrGrGrGrGrCrCrUrUrGrGrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrArUrGrUrArGrArCrArArGrGrCrCrCrCrCrC*mA*mC*mC	240
PRNP +72 nick sgRNA	mG*mC*mA*rUrGrUrUrUrUrCrArCrGrArUrArGrUrArArGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrmUrGrCrUmU*mU*mU	241
FANCF +5 G to T pegRNA	mG*mG*mA*rArUrCrCrCrUrUrCrUrGrCrArGrCrArCrCrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrGrGrArArArArArArGrUrGrGrCrUrGrCrArArGrGrUrCrArArGrGrUrGrUrGrCrUrGrCrArGrArArG*mG*mG*mA	242
FANCF +48 nick sgRNA	mG*mG*mG*rGrUrCrCrCrArGrGrUrGrCrUrGrArCrGrUrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrUmU*mU*mU	243
RNF2 +1 T insertion pegRNA	mG*mU*mC*rArUrCrUrUrArGrUrCrArUrUrArCrCrUrGrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrGrArArCrGrArArCrArCrUrCrArGrArGrUrArArUrGrArCrUrArArG*mA*mU*mG	244
RNF2 +41 nick sgRNA	mU*mC*mA*rArCrCrArUrUrArArGrCrArArArArCrArUrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrUmU*mU*mU	245
CXCR4 P191A install pegRNA	mC*mA*mA*rCrCrArCrCrCrArCrArArGrUrCrArUrUrGrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrUrGrArCrCrGrCrUrUrCrUrArCrGrCrCrArArUrGrArCrUrUrGrUrGrGrU*mG*mG*mU	246

CXCR4 +43 nick sgRNA	mC*mA*mU*rCrUrUrUrGrCrCrArArCrGrUrCrArGrUrGrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrCrGrUrUrArUrCrArArCrUrUrGrArArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrUmU*mU*mU	247
IL2RB H134D Y135F install pegRNA	mC*mC*mA*rGrGrUrGrUrCrUrUrUrCrArArArGrUrArGrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrUrCrCrCrArArGrCrCrUrCrGrArCrUrUrCrUrUrUrGrArArArGrA*mC*mA*mC	248
IL2RB +55 nick sgRNA	mC*mU*mC*rCrCrUrCrCrArArGrUrUrGrUrCrCrArCrGrGrUrUrUrUrArGrArGrCrUrArGrArArArUrArGrCrArArArGrUrUrArArArArUrArArGrGrCrUrArGrUrCrCrGrUrUrArUrCrArArCrUrUrGrArArArArArGrUrGrGrCrArCrCrGrArGrUrCrGrGrUrGrCrUmU*mU*mU	249

[735] Across these six disease-relevant edits, it was observed that PE4max increased average prime editing efficiency over PE2 by 29-fold in HeLa cells and 2.1-fold in HEK293T cells (FIG. 46D). Notably, PE4max editing efficiencies (8.6% editing with 0.19% indels in HeLa cells, and 20% editing with 0.26% indels in HEK293T cells) were similar to or exceeded those of PE3 (4.5% editing with 1.5% indels in HeLa cells, and 24% editing with 5.4% indels in HEK293T cells), but with far fewer indels. In addition, PE5max improved disease-relevant allele conversion over PE3 by an average of 6.1-fold in HeLa cells and 1.5-fold in HEK293T cells, and enhanced edit:indel purity by 6.4-fold in HeLa cells and by 3.5-fold in HEK293T cells (FIG. 46D). Taken together, these results demonstrate that PE4max and PE5max support substantially higher prime editing performance compared to PE2 and PE3 at gene targets relevant to human disease in commonly used cell lines.

[736] Next, PE4 and PE5 editing systems were evaluated in a cell model of genetic disease and in primary human cells. First, the pathogenic *CDKL5* c.1412delA mutation in human induced pluripotent stem cells (iPSCs) derived from a patient heterozygous for this allele was corrected (Chen *et al*., 2021). Electroporation of these iPSCs with PE3 components (*in vitro*-transcribed PE2 mRNA and synthetic pegRNAs and nicking sgRNAs) yielded 17% correction of editable pathogenic alleles and 20% total indel products (FIG. 46E, FIG. 56F). Co-electroporation of these components with MLH1dn mRNA for PE5 editing elevated correction efficiency to 34% and lowered the frequency of indels to 6.1%. To further minimize indels, PE4 and PE5b strategies were also used. Without complementary-strand nicking, it was observed that MLH1dn improved allele correction from 4.0% (PE2) to 10% (PE4) with few indels (<0.34%) (FIG. 46E, FIG. 56F). Similarly, PE3b resulted in 13% editing of the mutant allele with 4.8% indels, while PE5b elevated editing to 27% with 3.8% indels. Across the PE4, PE5, and PE5b systems tested,

MLH1dn enhances correction of the pathogenic *CDKL5* c.1412delA mutation by 2.2-fold in efficiency and 3.6-fold in outcome purity in patient-derived iPSCs.

[737] Second, primary T cells were electroporated from healthy human donors with PE2 mRNA, MLH1dn mRNA, and synthetic pegRNAs and nicking sgRNAs to introduce the protective *PRNP* G127V mutation, a G•C-to-T•A transversion at *FANCF*, or a 1-bp insertion of a T at *RNF2*. At these three sites, it was found that PE5 achieved an average of 41% editing efficiency with 12% indels, a substantial improvement from 22% editing efficiency and 26% indels with PE3 (FIG. 46F, FIG. 56G). PE4 and PE5 were also applied to install the protective *CXCR4* P191A allele that prevents HIV infection (Liu *et al.*, 2018), and the *IL2RB* H134D Y135F variant that enables orthogonal IL-2 T cell stimulation (Sockolosky *et al.*, 2018). For these edits, PE4 increased the frequency of allele conversion over PE2 by an average of 3.6-fold with few indel byproducts (FIG. 46F, FIG. 56G). In addition, PE5 resulted in an average of 52% editing with 11% indels, compared to 44% editing with 17% indels with PE3, a 2.0-fold improvement in edit:indel ratio. Taken together, these results across six sites in human iPSCs and primary T cells establish PE4 and PE5 as enhanced prime editing systems that enable substantially greater editing efficiency and outcome purity in cell types relevant to the study and potential treatment of genetic disease.

Discussion

[738] Using pooled CRISPRi screens, it was discovered that MMR activity strongly suppresses the efficiency and outcome purity of substitution prime edits. Based on the results and insights from the CRISPRi screens and the role of MMR in suppressing editing with nucleotide substitution, the PE4 and PE5 systems were developed. Particularly, the PE4 and PE5 systems co-express MLH1dn to transiently inhibit MMR, enhance prime editing efficacy, and reduce indels without inducing substantial off-target genomic changes. Optimization of the prime editor protein resulted in a PEmax architecture that can synergize with PE4 and PE5 systems, as well as with epegRNAs (Nelson *et al.*, 2021 - incorporated herein by reference), to further enhance prime editing performance. Together, the model for DNA repair of prime editing supported by these findings, the PE4 and PE5 strategies developed to circumvent a prime editing bottleneck, and the improved PEmax prime editor architecture described here substantially advance the utility of prime editing for precision manipulation of the genome.

[739] This work revealed that prime editing can install certain types of edits, including G•C-to-C•G transversions and substitutions of three or more contiguous bases, with higher efficiency due to the ability of the corresponding prime editing intermediates to evade MMR. In addition to edit type, other properties could also affect the sensitivity of prime editing to MMR, such as the sequence context of the target site. The state of the edited locus may also be important, as MMR more efficiently repairs early replicating euchromatin (Supek and Lehner, 2015) as well as lagging strand DNA during replication (Lujan *et al.*, 2014).

[740] While it is demonstrated that co-delivery of MLH1dn with prime editors enhances editing efficiency and precision, the studies presented herein on the role of MMR in determining prime editing outcomes coupled with insights into the types of prime editing intermediates that are repaired by MMR allows researchers to design prime editing experiments to evade MMR, even without expression of MLH1dn. For example, the data presented herein show that strategically installing additional nearby silent mutations can enhance prime editing outcomes by avoiding MMR reversal of prime editing intermediates, even with PE2 or PE3 systems. Other modalities for MMR inhibition may also prove beneficial for prime editing. Although no small molecules that selectively target MMR have been reported, chemical inhibitors would be useful in applications limited by MLH1dn delivery and could allow temporal control of MMR inhibition. For uses such as viral delivery that may maintain prime editor expression over a long duration, RNA interference offers another strategy for transient MMR knockdown.

[741] PE4 and PE5 systems powerfully enhance prime editing in seven mammalian cell types tested, including patient-derived iPSCs and primary T cells. PE4 increases editing efficiency by an average of 7.7-fold over PE2 in MMR-proficient cell types, uniquely enabling appreciable levels of gene editing without generating DSBs from complementary-strand nicking. This improvement in combination with the PEmax architecture and epegRNAs can surpass editing frequencies of PE3 while maintaining few indel byproducts. PE4max with epegRNAs thus will be particularly useful for gene editing applications that cannot tolerate indel formation or are limited by delivery of nicking sgRNAs. In comparison, PE5 elevates editing efficiency over PE3 by an average 2.0-fold and editing outcome purity by approximately 3-fold in cells with active MMR, which includes most cell targets.

Experimental Model and Subject Details

Culture conditions for immortalized cell lines

[742] HEK293T, HeLa dCas9–BFP–KRAB, HeLa, HCT116, and N2A cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) plus GlutaMAX (Thermo Fisher Scientific) supplemented with 10% fetal bovine serum (FBS) (Thermo Fisher Scientific). HeLa dCas9–BFP–KRAB cells were cultured in DMEM plus GlutaMAX supplemented with 10% FBS, 100 U mL-1 penicillin, and 100 µg mL-1 streptomycin (Thermo Fisher Scientific). K562 dCas9–BFP–KRAB and K562 cells were cultured in Roswell Park Memorial Institute (RPMI) 1640 medium (Thermo Fisher Scientific) supplemented with 10% FBS and 1% Penicillin-Streptomycin-Glutamine 100× (Thermo Fisher Scientific). All HAP1 cell types were cultured in Iscove's Modified Dulbecco's Medium (IMDM) plus GlutaMAX (Thermo Fisher Scientific) supplemented with 10% FBS. U2OS cells were cultured in McCoy's 5A medium (Gibco) supplemented with 10% FBS, 100 U mL^{-1} penicillin, and 100 µg mL^{-1} streptomycin (Thermo Fisher Scientific). HeLa dCas9–BFP–KRAB and K562 dCas9–BFP–KRAB cell lines were verified by short tandem repeat marker testing. All cell types were passaged every 2–3 days, maintained below 80% confluency, cultured at 37 °C with 5% CO_2, and tested negative for mycoplasma.

Isolation of primary human T cells

[743] Peripheral blood mononuclear cells (PBMCs) were isolated from buffy coat of healthy donors (Memorial Blood Centers in St. Paul, Minnesota) by density centrifugation using Lymphoprep density gradient medium (STEMCELL Technologies) and SepMate tubes (STEMCELL Technologies). T cells were isolated from PBMCs using the EasySep Human T Cell Isolation Kit (STEMCELL Technologies).

Culture conditions for human patient-derived induced pluripotent stem cells

[744] All iPSC culturing work was performed by staff at the Human Neuron Core at Boston Children's Hospital following institutional guidelines and under institutional approvals (IRB#: P00016119). A clonal iPS cell line, MAN0855-01 #A (Coriell Institute #OR00007), was expanded from a female *CDKL5* deficiency disorder patient carrying a heterozygous *CDKL5* c.1412delA p.D471fs mutation on the X chromosome (Chen *et al.*, 2021). MAN0855-01 #A was previously verified to express the mutant *CDKL5* transcript by Sanger sequencing of cDNA. The

MAN0855-01 #A iPS cell line was cultured in StemFlex medium (Thermo Fisher Scientific) on Geltrex (Thermo Fisher Scientific) diluted 1:50 in DMEM/F12 (Thermo Fisher Scientific) and coated according to the manufacturer's protocol. For regular maintenance, iPS cell colonies were clump-passaged using Gentle Cell Dissociation Reagent (STEMCELL Technologies) at 80% confluency every 5–7 days.

Method Details

General methods and molecular cloning

[745] Lentiviral transfer plasmids and plasmids for mammalian expression of prime editors and other proteins were cloned using uracil excision (USER) assembly (Cavaleiro *et al.*, 2015). Briefly, DNA fragments were amplified with deoxyuracil-containing primers (Integrated DNA Technologies) using the uracil tolerant Phusion U Green Multiplex PCR Master Mix (Thermo Fisher Scientific). Deoxyuracil-incorporated DNA fragments were assembled with USER enzyme (New England BioLabs) and DpnI (New England BioLabs) according to the manufacturer's protocol using junctions with a melting temperature of 42–60 °C, followed by transformation into cells. All prime editor constructs were cloned into the pCMV-PE2 vector backbone (Anzalone *et al.*, 2019) (Addgene #132775) under constitutive expression from a CMV promoter. All prime editor constructs also contained the following mutations within the MMLV RT: D200N, T306K, W313F, T330P, and L603W. All DNA repair protein and RFP expression constructs were cloned into vectors under constitutive expression from an EF1α promoter. Human MSH2, MSH6, PMS2, and MLH1 sequences were subcloned from the plasmids pFB1_hMSH2 (Addgene #129423), pFB1_hMSH6 (Addgene #129424), pFB1_PMS2 (Addgene #129425), and pFB1_MLH1 (Addgene #129426) (Geng *et al.*, 2011). Human CDKN1A sequence was subcloned from the plasmid Flag p21 WT (Addgene #16240) (Zhou *et al.*, 2001). Codon-optimized MLH1 sequences for human cell and mouse cell expression were designed using GenSmart Codon Optimization (Genscript) and ordered as gBlock gene fragments (Integrated DNA Technologies).

[746] Plasmids for mammalian expression of pegRNAs or sgRNAs were cloned using Golden Gate assembly (Engler *et al.*, 2008) as previously described (Anzalone *et al.*, 2019). Briefly, a guide RNA vector backbone for human U6 promoter expression was digested overnight with BsaI-HFv2 (New England BioLabs) according to the manufacturer's protocol, and linearized

product was purified by electrophoresis with a 1% agarose gel using the QIAquick Gel Extraction Kit (QIAGEN). Oligonucleotides (Integrated DNA Technologies) for the spacer sequence, guide RNA scaffold, and 3′ extension were annealed, assembled with linearized U6 backbone DNA using T4 DNA ligase (New England BioLabs) according to the manufacturer's protocol, and transformed into cells. Only guide RNA scaffold oligonucleotides were purchased with 5′ phosphorylation modifications. Some plasmids encoding pegRNAs and epegRNAs were synthesized by Twist Bioscience. A list of pegRNAs and nicking sgRNAs used in this work is provided in Table 7.

[747] Unless otherwise noted, assembled plasmids were transformed into One Shot Mach1 cells (Thermo Fisher Scientific) and grown on Luria-Bertani (LB) or 2×YT agar with 50 μg ml^{-1} carbenicillin (Gold Biotechnology). Plasmid sequences were fully verified by Sanger sequencing (Quintara Biosciences), and bacteria containing verified plasmids were grown in 2×YT medium with 100 μg ml^{-1} carbenicillin (Gold Biotechnology). Plasmid DNA were isolated using the QIAGEN Plasmid Plus Midi Kit or QIAGEN Plasmid Plus Maxi Kit with endotoxin removal and 2× the recommended amount of RNase A in Buffer P1. Some pegRNA and sgRNA plasmid DNA were isolated with the PureYield Plasmid Miniprep System (Promega Corporation) with endotoxin removal. Plasmid DNA purified using the PureYield Plasmid Miniprep System were only used for HEK293T and HeLa cell transfections. All plasmids were eluted in nuclease-free water (QIAGEN) and quantified using a NanoDrop One UV-Vis spectrophotometer (Thermo Fisher Scientific).

Lentivirus production for generating cell lines

[748] To package lentivirus for generating stable cell lines, HEK293T cells were seeded on 6-well plates (Corning) at 7.5×10^5 cells per well in DMEM supplemented with 10% FBS. At 60% confluency 16 h after seeding, cells were transfected with 12 μL Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol and 1.33 μg lentiviral transfer plasmid, 0.67 μg pMD2.G (Addgene #12259), and 1 μg psPAX2 (Addgene #12260). 6 h after transfection, media was exchanged with DMEM supplemented with 10% FBS. 48 h after transfection, viral supernatant was centrifuged at 3000 *g* for 15 min to remove cellular debris, filtered through a 0.45 μm PVDF filter (Corning), and stored at -80 °C.

Design and construction of HEK293T line with integrated HBB coding region

[749] A lentiviral transfer plasmid was designed and cloned to contain the human *HBB* coding region (CDS) and a PuroR–T2A–BFP marker under expression from an EF1α promoter (pEF1α). Lentivirus carrying this cassette were produced from HEK293T cells as described above. To stably integrate the *HBB* CDS, 6×10^5 HEK293T cells were infected with lentivirus in 6-well plates (Corning) with DMEM supplemented with 10% FBS and 10 μg mL^{-1} polybrene (Sigma-Aldrich). BFP fluorescence was monitored daily using a CytoFLEX S Flow Cytometer (Beckman Coulter) to ensure an MOI of 0.1 and single copy integration. Following infection for 2 days, HEK293T cells were selected in 2 μg μL^{-1} puromycin (Thermo Fisher Scientific) for 3 days and stable transduction was confirmed by measuring BFP fluorescence. The resulting cell line was used to optimize pegRNAs for prime editing at an integrated *HBB* site with SaPE2. To measure editing, a 214-bp amplicon of the integrated *HBB* CDS locus was PCR amplified. Amplification of the genomic *HBB* locus with these primers yields a differently sized 1064-bp amplicon.

Design and construction of HeLa line with CRISPRi sgRNA and prime edit target

[750] The lentiviral transfer plasmid backbone for prime editing Repair-seq screens (pPC1000) was designed and cloned to contain a universal prime edit site and express a control *S. pyogenes* sgRNA for CRISPRi targeting. The prime edit site consisted of an *HBB* target protospacer for Sa-pegRNA that is flanked by two complementary-strand Sa-sgRNA protospacers derived from the *Saccharomyces cerevisiae* genome. These protospacers were situated such that SaPE2–sgRNA complexes nick 50-bp upstream and 50-bp downstream of the nick formed by SaPE2–pegRNA. This 234-bp edit site was positioned adjacently to the sequence of an EGFP-targeting control *S. pyogenes* sgRNA under expression from a mouse U6 promoter such that sgRNA and edit site could be amplified by PCR in the same 453-bp amplicon. pPC1000 also contained an pEF1α–PuroR–T2A–BFP selection marker.

[751] Lentivirus encoding the pPC1000 cassette were produced from HEK293T cells as described above. For stable integration, 2.5×10^5 HeLa dCas9–BFP–KRAB cells were infected with pPC1000 lentivirus in 6-well plates (Corning) with DMEM supplemented with 10% FBS and 10 μg mL^{-1} polybrene (Sigma-Aldrich). BFP fluorescence was monitored with a CytoFLEX S Flow Cytometer (Beckman Coulter) to ensure an MOI of 0.1 and single copy integration.

Following 2 days of infection, cells were selected in 2 μg μL^{-1} puromycin (Thermo Fisher Scientific) for 3 days and stable transduction was confirmed by measuring BFP fluorescence. The resulting HeLa dCas9–BFP–KRAB cell line with integrated pPC1000 sequence was used to pilot prime editing conditions, Sa-pegRNAs, and Sa-sgRNAs for Repair-seq screens.

Transfection of HEK293T, HeLa, HCT116, and N2A cells

[752] Unless otherwise noted, HEK293T cells were seeded on 96-well plates (Corning) at 1.6–1.8 × 10^4 cells per well in DMEM plus GlutaMAX supplemented with 10% FBS. Between 16 and 24 h after seeding, cells were transfected at 60–80% confluency with 0.5 μL Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol and 200 ng prime editor plasmid, 66 ng pegRNA plasmid, 22 ng sgRNA plasmid (when required), and 100 ng for RFP or MMR protein expression (when required).

[753] For arrayed experiments, HeLa dCas9–BFP–KRAB and HeLa cells were seeded on 96-well plates (Corning) at 8 × 10^3 cells per well in DMEM plus GlutaMAX supplemented with 10% FBS. Between 16 and 24 h after seeding, cells were transfected at 60–80% confluency with 0.3 μL *Trans*IT-HeLa reagent (Mirus Bio) according to the manufacturer's protocol and 56.25 ng prime editor plasmid bearing a P2A–BlastR selection marker, 18.75 ng pegRNA plasmid, 6.25 sgRNA plasmid (when required), and 28.1 ng human codon-optimized MLH1dn plasmid (when required). 24 h following transfection, 10 ng μL^{-1} blasticidin (Thermo Fisher Scientific) was added to each well to select for cells expressing prime editor.

[754] HCT116 cells were seeded on 96-well plates (Corning) at 1.6 × 10^4 cells per well in DMEM plus GlutaMAX supplemented with 10% FBS. Between 16 and 20 h after seeding, cells were transfected at 60–80% confluency with 0.5 μL Lipofectamine 3000 plus 0.8 μL P3000 reagent (Thermo Fisher Scientific) according to the manufacturer's protocol and 200 ng prime editor plasmid bearing a P2A–BlastR selection marker, 66 ng pegRNA plasmid, 22 ng sgRNA plasmid (when required), and 100 ng MLH1dn plasmid (when required). The day after transfection, media was replaced with fresh DMEM plus GlutaMAX supplemented with 10% FBS and 10 ng μL^{-1} blasticidin (Thermo Fisher Scientific) to select for cells expressing prime editor.

[755] N2A cells were seeded on 96-well plates (Corning) at 1.6 × 10^4 cells per well in DMEM plus GlutaMAX supplemented with 10% FBS. Between 16 and 20 h after seeding, cells were

transfected at 60–80% confluency with 0.5 µL Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol and 175 ng prime editor plasmid, 50 ng pegRNA plasmid, 20 ng sgRNA plasmid (when required), and 87.5 ng plasmid encoding human codon-optimized hMLH1dn or mouse codon-optimized mMLH1dn when required. Genomic DNA was extracted 72 h following transfection.

Electroporation of HAP1, K562, and U2OS cells

[756] HAP1 cells were nucleofected using the SE Cell Line 4D-Nucleofector X Kit S (Lonza) according to the manufacturer's protocol with 4×10^5 cells (program DZ-113), 300 ng PE2–P2A–BSD, 100 ng pegRNA plasmid, and 33 ng sgRNA plasmid (when required). After nucleofection, cells were cultured in 48-well plates (Corning) with IMDM plus GlutaMAX supplemented with 10% FBS. The day after nucleofection, media was replaced with fresh IMDM plus GlutaMAX supplemented with 10% FBS and 10 ng μL^{-1} blasticidin (Thermo Fisher Scientific) to select for cells expressing prime editor.

[757] K562 cells were nucleofected using the SF Cell Line 4D-Nucleofector X Kit S (Lonza) according to the manufacturer's protocol with 5×10^5 cells (program FF-120), 800 ng prime editor plasmid, 200 ng pegRNA plasmid, 83 ng sgRNA plasmid (when required), and 400 ng MLH1dn plasmid (when required). After nucleofection, cells were cultured in 6-well plates (Corning) with RPMI 1640 medium supplemented with 10% FBS and 292 µg mL^{-1} L-Glutamine (Thermo Fisher Scientific).

[758] U2OS cells were nucleofected using the SE Cell Line 4D-Nucleofector X Kit S (Lonza) according to the manufacturer's protocol with 2×10^5 cells (program DN-100), 1600 ng PE2 or PE2–P2A–MLH1dn plasmid, 400 ng pegRNA plasmid, and 166 ng sgRNA plasmid (when required). After nucleofection, cells were cultured in 12-well or 24-well plates (Corning) with McCoy's 5A medium supplemented with 10% FBS.

Genomic DNA extraction

[759] Unless otherwise noted, HEK293T, HeLa dCas9–BFP–KRAB, HeLa, HCT116, N2A, HAP1, K562, and U2OS cells were cultured for 72 h after transfection or nucleofection before genomic DNA was isolated. Cells were washed once with PBS (Thermo Fisher Scientific) and lysed with gDNA lysis buffer (10 mM Tris-HCl, pH 8.0; 0.05% SDS; 800 units μL^{-1} proteinase

K (New England BioLabs)) at 37 °C for 1.5–2 h, followed by enzyme inactivation at 80 °C for 30 min.

High-throughput amplicon sequencing of genomic DNA samples

[760] To assess gene editing, loci were amplified from genomic DNA samples via two rounds of PCR and deep sequenced. Briefly, an initial PCR step (PCR1) amplified the genomic sequence of interest using primers (Integrated DNA Technologies) containing Illumina forward and reverse adapters. Each 20 µL PCR1 reaction was performed with 500 nM of each primer, 0.8 to 1.0 µL genomic DNA, 1× SYBR Green (Thermo Fisher Scientific), and 10 µL Q5 High-Fidelity 2× Master Mix (New England BioLabs) on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories) with the following thermocycling conditions: 98 °C for 2 min, 29–31 cycles of [98 °C for 10 s, 61 °C for 20 s, and 72 °C for 30 s], followed by 72 °C for 2 min. PCR1 reactions were monitored with SYBR Green fluorescence to avoid over-amplification. A list of primers used for PCR1 reactions is provided in Table 8.

[761] Table 8. Primers Used for PCR1 Reactions

Primer description	**Primer sequence (5'-3')**	**SEQ ID NO:**
Lentiviral HBB CDS site - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAAATTGGACAGCAAGAAAGCGAGCTTAG	250
Lentiviral HBB CDS site - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTTTGCCACACTGAGTGAGCTG	251
Lentiviral pPC1000 edit site - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCCTTTAGTGATGGCCTGGCTCACCTG	252
Lentiviral pPC1000 edit site - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTTAGGCGGCCGCTGCACGTAG	253
Lentiviral pPC1000 CRISPRi sgRNA and edit site (for Repair-seq screens) - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTATCCCTTGGAGAACCACCTTGTTG	254
Lentiviral pPC1000 CRISPRi sgRNA and edit site (for Repair-seq screens) - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCAACCAACGCTATCGGCATGG	255
HEK3 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAGGGAAACGCCCATGCAATTAG	256
HEK3 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTCTGTTGAGCTCGACCCTGAA	257
EMX1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAGCTCAGCCTGAGTGTTGA	258
EMX1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTCGTGGGTTTGTGGTTGC	259
RNF2 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTAGCCAACATACAGAAGTCAGG	260
RNF2 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTAGACCATAGCACTTCCCTTCC	261

Primer description	Primer sequence (5'-3')	SEQ ID NO:
FANCF locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCGATGGATGTGGCGCAGGTAG	262
FANCF locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTAGACGCTGGGAGATTGACATGC	263
RUNX1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCACAAACAAGACAGGGAACTG	264
RUNX1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTAGATGTAGGGCTAGAGGGGTG	265
DNMT1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCACAACAGCTTCATGTCAGCC	266
DNMT1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTACGTTAATGTTTCCTGATGGTCC	267
VEGFA locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACTTGGTGCCAAATTCTTCTCC	268
VEGFA locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTAAAGAGGGAATGGGCTTTGGA	269
HEK4 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGAACCCAGGTAGCCAGAGAC	270
HEK4 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTTCCTTTCAACCCGAACGGAG	271
HBB locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCCTTTAGTGATGGCCTGGCTCACCTG	252
HBB locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTTCTCTGTCTCCACATGCC	272
PRNP locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTCAGTGGAACAAGCCGAGT	273
PRNP locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTACTTGGTTGGGGTAACGGTG	274
CDKL5 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTTCAGAAGGCCCAGGGACAAAG	275
CDKL5 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGGTACTGGGCTCCGCAATTT	276
CXCR4 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCTTGAGGGCCTTGCGCTTC	277
CXCR4 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCTTCATCAGTCTGGACCGCTA	278
IL2RB locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGACCAGGACAGGAAGGAGGAA	279
IL2RB locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTGGCAGGTACAAAGTGGGAGG	280
HBG1/2 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCTTCATCCCTAGCCAGCCGC	281
HBG1/2 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTAACGGCTGACAAAAGAAGTCCTGG	282
mouse Dnmt1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCCTTCGGGCATAGCATGG	283
mouse Dnmt1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCTCGGCATCGGTCC	284
mouse Chd2 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAGTCAGTGCCTCACCTCTCACAC	285

Primer description	Primer sequence (5'-3')	SEQ ID NO:
mouse Chd2 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTTGCAGATCGAGGAGACTGGCA	286
mouse Col12a1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGAAGTCATGTGCGGTCTGGTCA	287
mouse Col12a1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTAGCTCACAGCCTCATGAAGGGA	288
mouse Ctnnb1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACCAGCTACTTGCTCTTGCGTG	289
mouse Ctnnb1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCGTAGATGGCTTCTTCAGGTAGCA	290
HEK3 OT1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCCCCTGTTGACCTGGAGAA	291
HEK3 OT1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGTACTTGCCCTGACCA	292
HEK3 OT2 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTGGTGTTGACAGGGAGCAA	293
HEK3 OT2 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGAGATGTGGGCAGAAGGG	294
HEK3 OT3 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTGAGAGGGAACAGAAGGGCT	295
HEK3 OT3 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCCAAAGGCCCAAGAACCT	296
HEK3 OT4 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTCCTAGCACTTTGGAAGGTCG	297
HEK3 OT4 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCTCATCTTAATCTGCTCAGCC	298
FANCF OT1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCGGGCAGTGGCGTCTTAGTCG	299
FANCF OT1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCCTGGGTTTGGTTGGCTGCTC	300
FANCF OT2 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTCCTTGCCGCCCAGCCGGTC	301
FANCF OT2 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTGGGGAAGAGGCGAGGACAC	302
FANCF OT3 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCCAGTGTTTCCCATCCCCAACAC	303
FANCF OT3 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAATGGATCCCCCCCTAGAGCTC	304
FANCF OT4 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAGGCCCACAGGTCCTTCTGGA	305
FANCF OT4 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCACACGGAAGGCTGACCACG	306
EMX1 OT1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGTGGGGAGATTTGCATCTGTGGAGG	307
EMX1 OT1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCTTTTATACCATCTTGGGGTTACAG	308
EMX1 OT2 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCAATGTGCTTCAACCCATCACGGC	309
EMX1 OT2 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCATGAATTTGTGATGGATGCAGTCTG	310

Primer description	Primer sequence (5'-3')	SEQ ID NO:
EMX1 OT3 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGAGAAGGAGGTGCAGGAGCTAGAC	311
EMX1 OT3 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCATCCCGACCTTCATCCCTCCTGG	312
EMX1 OT4 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNACAGGTGAATAACCGGTCGCCA	313
EMX1 OT4 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTTTTATCAAACAAGGTGCAGATACAGCAA	314
HEK4 OT1 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGCATGGCTTCTGAGACTCA	315
HEK4 OT1 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGTCTCCCTTGCACTCCCTGTCTTT	316
HEK4 OT2 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTTGGCAATGGAGGCATTGG	317
HEK4 OT2 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGAAGAGGCTGCCCATGAGAG	318
HEK4 OT3 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGGTCTGAGGCTCGAATCCTG	319
HEK4 OT3 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGTGGCCTCCATATCCCTG	320
HEK4 OT4 locus - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNTTTCCACCAGAACTCAGCCC	321
HEK4 OT4 locus - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCTCGGTTCCTCCACAACAC	322
Bat-25 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGGAGGATGACGAGTTGGCCCTAGAC	323
Bat-25 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCAAAGAGACAGCAGTTGGAACATGA	324
Bat-26 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGTGGAGTGGAGGAGGGGAGAGAAA	325
Bat-26 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTTTGCAGTTTCATCACTGTCTGCGGT	326
MONO-27 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTACTGTCCTACTGTGCCTGGCTCC	327
MONO-27 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCAGCCTGGGCAAGATAATGAGACCC	328
NR-21 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGGTGCACAGAGCAGAACCATCCT	329
NR-21 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCAACCTCAAAAGCTGCCTCCCTTT	330
NR-24 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCTGTAGTCCCAGCTATTCGGGAGGC	331
NR-24 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTAAATGACCCCTTCCTGCCCATCACT	332
HSPH1-T17 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTGGAAAAGGAACTGCATCTGTGACGG	333
HSPH1-T17 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTTTTTCCTAATCCCCTGTGAAACCTGT	334
EWSR1 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAACAATGTTCATGGTTGTGATGT	335

Primer description	Primer sequence (5'-3')	SEQ ID NO:
EWSR1 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCTGGATCTC CCCTTGAACCTTTGGA	336
MSI-01 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTTGAT GTCCTGCGTCTAGGGTCTGC	337
MSI-01 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGACTGGAG CCTTGGACAGGTTGAGA	338
MSI-03 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCCAC TGCTATTTGAAAGAGTTGCTC	339
MSI-03 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACACCTGG CTAAATGCTCGGATT	340
MSI-04 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTTCCAA GATTCCTTCCCTGGCCACTC	341
MSI-04 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTACTGTCTGT AGTCCTGGCTTCGTGG	342
MSI-06 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCAG CAAACTGAACAGGTCACCAAC	343
MSI-06 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTGCTACACTG TGGCATCAGCACATATC	344
MSI-07 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTGA AAGCAACCTAAGCTGTGGTGA	345
MSI-07 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTTGCTATAAG AGCTGAGCAGACGACA	346
MSI-08 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTCCAGC CCCCATGTACACTGTAGTCG	347
MSI-08 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCCACCCCA AGGCCAAAATCAGTAA	348
MSI-09 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGTCTC GGCTACTTGGGAGGCTTAGG	349
MSI-09 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCGACTAAA GAGGTCATTCACTTGT	350
MSI-11 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTAGCAT GTTTGCAGCCTTCTTCTGGA	351
MSI-11 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTAGAACTTCT CTCACAATGTAGCCCCT	352
MSI-12 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTGA GGCTAAACACTATCATGCCA	353
MSI-12 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTCAGAGGTTG CAGTGAGCCGAGATTG	354
MSI-13 microsatellite - PCR1 forward	ACACTCTTTCCCTACACGACGCTCTTCCGATCTACATC TTCAGGTCAGGAAAACAGCTCG	355
MSI-13 microsatellite - PCR1 reverse	TGGAGTTCAGACGTGTGCTCTTCCGATCTTAATGACTT GGGCTTTGGAAGCAGC	356
MSH2 - RT-qPCR forward	CTTCGTGCGCTTCTTTCAGGG	357
MSH2 - RT-qPCR reverse	GCAGATTCTTTGCTCCTGCCG	358
MSH6 - RT-qPCR forward	CGCTGAGTGATGCCAACAAGG	359
MSH6 - RT-qPCR reverse	AGCCCTCCGTTGAGGTTCTTC	360
MLH1 - RT-qPCR forward	AAGACAATGGCACCGGGATCA	361
MLH1 - RT-qPCR reverse	AAGCCTCACCTCGAAAGCCAT	362
PMS2 - RT-qPCR forward	CTGGGCAGGTGGTACTGAGTC	363
PMS2 - RT-qPCR reverse	TAGTGGCACCAGCATCCAGAC	364
ACTB - RT-qPCR forward	GAGGCACTCTTCCAGCCTT	365

Primer description	Primer sequence (5'-3')	SEQ ID NO:
ACTB - RT-qPCR reverse	AAGGTAGTTTCGTGGATGCC	366

[762] The subsequent PCR step (PCR2) added unique i7 and i5 Illumina barcode combinations to both ends of the PCR1 DNA fragment to enable sample demultiplexing. Each 12.5 µL PCR2 reaction was performed with 500 nM of each barcoding primer, 0.5 µL PCR1 product, and 6.25 µL Phusion U Green Multiplex PCR Master Mix (Thermo Fisher Scientific) with the following thermocycling conditions: 98 °C for 2 min, 9 cycles of [98 °C for 15 s, 61 °C for 20 s, and 72 °C for 30 s], followed by 72 °C for 2 min. PCR2 products were pooled by common amplicons, separated by electrophoresis on a 1% agarose gel, purified using the QIAquick Gel Extraction Kit (QIAGEN), and eluted in nuclease-free water. DNA amplicon libraries were quantified with a Qubit 3.0 Fluorometer (Thermo Fisher Scientific), then sequenced using the MiSeq Reagent Kit v2 or MiSeq Reagent Micro Kit v2 (Illumina), with 280–300 single-read cycles. A list of FASTQ sequencing files generated in this work is provided in Table 7.

Quantification of amplicon sequencing data

[763] All arrayed prime editing experiments from FIGs. 42A-54B were analyzed as follows. Sequencing reads were demultiplexed using MiSeq Reporter (Illumina). Amplicon sequences were aligned to a reference sequence with CRISPResso2 (Clement *et al.*, 2019) in standard mode using the parameters "-q 30" and "-discard_indel_reads TRUE". For each amplicon, the CRISPResso2 quantification window was positioned to include the entire sequence between pegRNA- and sgRNA-directed Cas9 cut sites, as well as an additional ≥10-nt beyond both cut sites. For each amplicon, the same quantification window was used for PE2, PE3, PE4, and PE5 conditions, regardless of whether a nicking sgRNA is used. All prime editing efficiencies describe percentage of (number of reads with the intended edit that do not contain indels)/(number of reads that align to the amplicon). Single-base substitution prime editing frequencies were quantified as: (frequency of intended base substitution in reference-aligned, non-discarded reads) × (number of reference-aligned, non-discarded reads)/(number of reference-aligned reads). For all other prime edits (insertion, deletion, contiguous substitutions, combinations of edits), CRISPResso2 was run in HDR mode with all the same parameters described above and using the intended editing outcome as the expected allele (-e). Frequencies for these edits were quantified as: (number of HDR-aligned reads)/(number of reference-aligned

reads). All indel frequencies were quantified as: (number of indel-containing reads)/(number of reference-aligned reads).

CRISPRi library cloning and lentivirus production

[764] An oligonucleotide library of CRISPRi sgRNAs, oBA697, was designed to contain 60 non-targeting control sgRNAs and 1,513 sgRNAs that target 477 DNA repair genes with three sgRNAs per gene (Hussmann *et al.*, 2021). A list of targeted genes and sequences in the sgRNA library is provided in Table 5. The oBA697 oligonucleotide library was amplified by PCR using Phusion High-Fidelity DNA Polymerase (New England BioLabs) and purified with the NucleoSpin Gel and PCR Clean-up kit (Macherey-Nagel). The amplified CRISPRi library inserts and the pPC1000 lentiviral screen vector containing the pre-validated prime edit site were digested with BstXI and BlpI restriction endonucleases (Thermo Fisher Scientific), ligated with T4 ligase (New England BioLabs), then transformed into Stellar cells (Takara Bio). The pooled plasmid library was amplified in MegaX DH10B electrocompetent cells (Thermo Fisher Scientific). The pooled library sequences were further verified by sequencing on a MiSeq Reagent Kit v2 (Illumina).

[765] To produce lentivirus with pPC1000 libraries, HEK293T cells were seeded in a 15 cm dish with DMEM supplemented with 10% FBS. One day after seeding, cells were transfected using 60 µL *Trans*IT-LT1 reagent (Mirus Bio) with 15 µg pPC1000 plasmid library and 5 µg packaging plasmids for expression of HIV-1 gag/pol, rev, tat, and VSV-G envelope protein. 24 h after transfection, 40 µL ViralBoost reagent (ALSTEM) was added to each 15 cm dish. 48 h after transfection, viral supernatant was collected, filtered through a 0.45 µm PVDF filter (Corning), and stored at -80 °C.

Repair-seq screens in HeLa cells

[766] PE2, PE3+50, PE3–50 Repair-seq screens were performed in biological duplicate in HeLa cells with integrated dCas9–BFP–KRAB (Addgene #46911) (Gilbert *et al.*, 2013). Lentivirus containing the CRISPRi library were produced from HEK293T cells as described above. HeLa CRISPRi cells expressing dCas9–KRAB were transduced with the lentiviral library at 0.1 MOI in DMEM supplemented with 10% FBS, 100 U mL^{-1} penicillin, 100 µg mL^{-1} streptomycin (Thermo Fisher Scientific), and 8 µg mL^{-1} polybrene (Sigma-Aldrich). 2 days after infection, HeLa CRISPRi cells were treated with 1 µg mL^{-1} puromycin (Thermo Fisher

Scientific) to select for HeLa CRISPRi cells with integrated library members. 3 days after infection, an additional 2 μg mL^{-1} puromycin was added to cells. Throughout lentiviral transduction and selection steps, cells were analyzed for BFP fluorescence on a BD LSRII flow cytometer to ensure a MOI of 0.1 and completed selection. Following 3 days of selection, media was changed to DMEM supplemented with 10% FBS, and HeLa CRISPRi cells were transfected at 50% confluency in 150 mm culture dishes (Corning) with 30 μg SaPE2–P2A–BSD plasmid, 10 μg Sa-pegRNA plasmid for installing a +6 G•C to C•G edit at the pre-validated edit site, and 3.3 μg Sa-sgRNA plasmid for +50 or –50 complementary-strand nicking (when required), using 140 μL *Trans*IT-HeLa reagent (Mirus Bio) according to the manufacturer's protocol. For an unedited HeLa control condition, cells were transfected with only 30 μg SaPE2–P2A–BSD plasmid as described above. 24 h following transfection, cells were treated with 10 μg mL^{-1} blasticidin (Thermo Fisher Scientific) to select for expression of SaPE2 protein. 72 h after transfection, HeLa CRISPRi cells were washed with PBS (Thermo Fisher Scientific), resuspended using Trypsin and DMEM, and pelleted at 1000 *g* for 10 min. Finally, cells were washed once more with PBS, pelleted at 1000 *g* for 10 min, then stored at -80°C.

Repair-seq screens in K562 cells

[767] PE2 and PE3+50 Repair-seq screens were performed in biological duplicate in K562 cells with integrated dCas9–BFP–KRAB (Addgene #46911) (Gilbert *et al.*, 2014). Lentivirus containing the CRISPRi library were produced from HEK293T cells as described above. K562 CRISPRi cells expressing dCas9-KRAB were transduced with the lentiviral library at 0.2 MOI in RPMI supplemented with 10% FBS, 100 U mL^{-1} penicillin, 100 μg mL^{-1} streptomycin, 292 μg mL^{-1} L-Glutamine, and 8 μg mL^{-1} polybrene (Sigma-Aldrich) by centrifugation at 1000 *g* for 2 h at room temperature. 2 days post infection, K562 CRISPRi cells were treated with 3 μg mL^{-1} puromycin (Thermo Fisher Scientific) to select for cells with integrated library members. After infection, the density of the K562 CRISPRi cells was maintained at approximately 5×10^5 mL^{-1} and the culture was replaced by fresh RPMI supplemented with 10% FBS, 100 U mL^{-1} penicillin, 100 μg mL^{-1} streptomycin, 292 μg mL^{-1} L-Glutamine, and 3 μg mL^{-1} puromycin 3 days and 5 days post infection. During media replacement, the cells were pelleted, washed with DPBS and resuspended in fresh media to remove dead cells. All centrifugations were performed at 150 *g* for 5 min in 50 mL canonical tubes to avoid cell loss. 6 days post infection, the culture was replaced

twice by fresh RPMI supplemented with 10% FBS and 292 μg mL^{-1} L-Glutamine (Thermo Fisher Scientific) to remove dead cells and antibiotics. Throughout lentiviral transduction and selection steps, cells were analyzed for BFP fluorescence on an Attune NxT flow cytometer to ensure an MOI of 0.2 and completed selection. 7 days post infection, the cells were nucleofected using the SE Cell Line 4D-Nucleofector X kit L (Lonza) with 1×10^7 cells (program FF-120), 7.5 μg SaPE2 plasmid, 2.5 μg Sa-pegRNA plasmid for installing a +6 G•C to C•G edit at the pre-validated edit site, and 3.3 μg Sa-sgRNA plasmid for +50 complementary-strand nicking (for PE3+50 conditions). For an unedited K562 control condition, cells were mock electroporated without any DNA plasmid as described above. After nucleofection, the cells were seeded at a density of 5×10^5 mL^{-1} in RPMI supplemented with 10% FBS and 292 μg mL^{-1} L-Glutamine. 48 h post nucleofection, the culture were pipetted up and down 5 times to prevent cells from clumping. 84 h post nucleofection, the cells were pelleted at 1000 *g* for 10 min, washed with DPBS, pelleted at 1000 *g* for 10 min, and then stored at -80 °C.

High-throughput sequencing of Repair-seq libraries

[768] Genomic DNA from edited HeLa CRISPRi and K562 CRISPRi cells was extracted using NucleoSpin Blood XL Maxi kit (Machery-Nagel). The number of live cells from which genomic DNA was extracted for each Repair-seq condition is listed in Table 9.

[769] Table 9. High-throughput Sequencing of Repair-seq Libraries

PE repair-seq screen condition	Number of live cells harvested for genomic DNA extraction (× 10e6)	Number of sequencing reads
K562 PE2 rep1	83.72	118658517
K562 PE2 rep2	118.69	188344211
K562 PE3+50 rep1	70.213	230946050
K562 PE3+50 rep2	90.965	206570163
HeLa PE2 rep1	107.85	420685828
HeLa PE2 rep2	73.5	171390008
HeLa PE3+50 rep1	105.6	229364033
HeLa PE3+50 rep2	95.4	290852565
HeLa PE3–50 rep1	93.15	234577915
HeLa PE3–50 rep2	96.9	150012222

[770] The genomic DNA was used as template for the initial round of PCR (PCR1) to amplify the region of interest. For each sample, a 100 μL PCR1 reaction was performed with 1 μL of 100 μM of each primer for amplifying pPC1000 sgRNA and edit site (Table 8), 10 μg of genomic DNA as template, 50 μL of NEBNext Ultra II Q5 Master Mix (New England BioLabs), and

molecular biology grade water (Corning) on a BioRad C1000 thermal cycler with the following thermocycler conditions: 98 °C for 30 s, 22 cycles of [98 °C for 10 s, 65 °C for 75 s], followed by 65 °C for 5 min. These amplification reactions were verified by TBE gel electrophoresis and ethidium bromide staining. The amplified products were purified using SPRIselect (Beckman Coulter) prior to the following round of PCR amplification (PCR2). Purified PCR1 amplicons were quantified using a high sensitivity DNA chip (Agilent Technologies) on an Agilent 2100 Bioanalyzer. PCR2 enabled indexing of the samples by the addition of i7 and i5 Illumina barcodes. For each 50 µL indexing PCR2 reaction: 10 ng of PCR1 amplicon was used as template along with 25 µL of KAPA HiFi HotStart ReadyMix (Roche Molecular Systems), 3 µL of 10 µM of each barcoding primer with the following thermocycler conditions: 95 °C for 3 min, 8 cycles of [98 °C for 20 s, 65 °C for 15 s, 72 °C for 15 s], followed by 72 °C for 1 min. The reactions were verified by TBE gel electrophoresis and ethidium bromide staining. PCR2 products from four PCR2 products were purified using SPRIselect and quantified on an Agilent 2100 Bioanalyzer prior to pooling. Repair-seq libraries were sequenced with the NovaSeq 6000 S1 Reagent Kit v1.5 (Illumina) with two 8-nt index reads, 44 cycles for R1 read, 263 cycles for R2 read. The number of sequencing reads acquired for each screen condition and replicate are listed in Table 9.

Processing of Repair-seq screen data

[771] Repair-seq screen data was processed using a modified version of the analysis approach described in (Hussmann *et al.*, 2021), with modifications made to accommodate the different library preparation strategy used in this Example (direction amplification of genomic DNA without ligation of UMIs before amplification) and the qualitatively different categories of repair outcomes empirically observed in prime editing data.

[772] Briefly, sequencing data for a batch of screens consists of 4 reads per spot: 2 8-nt index reads, a 44-nt R1 read of the CRISPRi sgRNA, and a 263-nt R2 read of the repair outcome. Reads from a batch of screens are demultiplexed into individual screens based on index reads. Within each screen, reads are demultiplexed into sets representing outcomes from cells receiving each individual CRISPRi sgRNA by comparing R1 sequences to a table of expected CRISPRi sgRNAs, allowing up to one mismatch between observed and expected sequences. Because direct amplification without UMIs does not allow consensus error correction of multiple reads of

each repair outcome, analysis must account for presence of errors in outcome sequences introduced by PCR or by sequencing to avoid interpreting such errors as genuine editing outcomes. As an initial triage, reads with less than 60% of base calls with a quality score greater than or equal to 30 were discarded.

[773] To categorize a repair outcome sequencing read that passed this quality filter, the outcome was first locally aligned to the screen vector, the pegRNA sequence, the human genome (hg19) and the bos taurus genome (bosTau7) to identify a comprehensive set of alignments between portions of the outcome sequence and any of these reference sequences. The set of local alignments identified was then pruned to a parsimonious set of alignments that explains as much of the read as possible using the minimum number of alignments using a greedy approach. The parsimonious alignments are then parsed through a decision tree that examines their configuration to assign the outcome to a category.

[774] Outcomes were classified as unedited if they consisted of a single alignment to the screen vector that did not contain the programmed SNV or any indels, with the exception of deletions of 1-nt that did not fall within 5-nt of a programmed nick, which were considered possible sequencing or PCR errors and were disregarded. Outcomes were classified as deletions if they consisted of two alignments to the screen vector that collectively covered the entire read but omitted a segment of the screen vector. Outcomes were classified as tandem duplications if they consisted of two or more alignments to the screen vector that collectively covered the entire read such that the portions of the screen vector covered by any two consecutive alignments on the read overlapped. Outcomes were classified as joining of pegRNA sequence at an unintended location if the set of parsimonious alignments included an alignment to the pegRNA such that the primer binding site (PBS) of the pegRNA was aligned to the same part of the read as the PBS in a screen vector alignment but the reverse transcription template (RTT) of the pegRNA was not aligned to the same part of the read as the RTT in a screen vector alignment. Note that in some such cases, the sequence produced is also consistent in theory with a multi-stage editing event consisting of an initial deletion or duplication that does not disrupt the PAM or protospacer followed by pegRNA-dependent editing of the resulting modified target sequence. Outcomes were classified as installation of additional edits from nearly matched scaffold sequence if the set of parsimonious alignments included an alignment to the pegRNA such that both the PBS and RTT were aligned to the same parts of the read as the PBS and RTT in a single alignment to the

screen vector that covered the whole read and that the pegRNA alignment contained fewer edits relative to the outcome than the screen vector alignment.

Quantification of CRISPRi-induced changes in outcome frequencies

[775] Following categorization of all outcomes for all CRISPRi sgRNAs, counts of each category for each sgRNA are collected into a matrix for downstream analysis, and the total frequency of each category across all outcomes from cells receiving non-targeting sgRNAs is calculated to establish unperturbed baseline frequencies. Because not all CRISPRi sgRNAs achieve high levels of knockdown, calculation of gene-level effects of CRISPRi sgRNAs on outcome categories must strike a balance between assigning increased confidence to phenotypes that are supported by multiple sgRNAs per gene without penalizing genes if not all sgRNAs targeting the gene have high activity. To do this, gene-level changes in outcome category frequencies in a given screen replicate (used in FIGs. 40E–40G, FIGs. 41E–41H, FIGs. 47A–47I) are calculated by first computing the log2 fold change in frequency of the category for every targeting sgRNA relative to the combined frequency across all non-targeting sgRNAs. For each gene, the gene-level log2 fold change is then taken to be the mean of these values for the two sgRNAs targeting the gene with the most extreme absolute values. To provide a qualitative estimate of the range of values produced in the absence of genuine signal by this process of selecting extreme values, the sixty non-targeting sgRNAs were randomly partitioned into 20 sets of 3 quasi-genes and the same process was applied to these quasi-genes.

Quantification of deletion boundaries

[776] To maximize signal to noise in calculation of position-specific profiles of deletion frequencies and in relative fraction of deletions that removed sequence far outside of programmed nicks, outcomes from all sgRNAs targeting each gene were grouped together. In each such group of outcomes, an array of counts for every position in the screen vector was initialized to 0. For each read classified as a deletion, an interval from the first position deleted to the last position deleted was incremented by 1 in the array of counts. The final array of counts was then divided by the total number of outcomes. Deletions flanked by microhomology result in sequence outcomes that are consistent with two or more degenerate pairs of deletion boundaries. For these deletions, the pair with the minimum values in the coordinate system of the screen vector was arbitrarily chosen. To prevent primer dimers or other non-specific amplification

products from being incorrectly identified as long deletions, apparent deletions for which the deleted region overlapped a window of 10-nt around either amplicon primer were excluded from calculation of deletion boundary statistics.

Design of recoded Sa-pegRNA scaffold

[777] In Repair-seq screens performed, an unintended editing outcome was observed in which additional edits are installed from nearly-matched scaffold sequence. This outcome contains a +17 T•A-to-C•G and +19 C•G insertion, in addition to the intended +6 G•C-to-C•G transversion. These unintended edits are consistent with incorporation of an extended 3′ DNA flap generated from reverse transcription of the Sa-pegRNA scaffold sequence into the genome. Because this extended 3′ flap shares 5-nt of homology (5′–GCCAA–3′) with the genomic target sequence after the last edited nucleotide, it was hypothesized that disrupting this homology could reduce the frequency of incorporating these unintended edits from reverse transcription of the Sa-pegRNA scaffold. A recoded Sa-pegRNA was therefore designed that alters two base pairs within the Sa-pegRNA scaffold while preserving the same base pairing interactions. The extended 3′ flap templated by this recoded Sa-pegRNA has reduced homology with the genomic target sequence. The spacer, PBS, and RT template sequences of the recoded Sa-pegRNA are identical to those for the Sa-pegRNA used in Repair-seq screens. It was observed that prime editing with this recoded Sa-pegRNA mediates similar frequencies of intended editing but substantially reduced unintended scaffold sequence incorporation compared to the original Sa-pegRNA used in Repair-seq screens.

HEK293T siRNA transfection

For experiments in **FIG. 42C**, HEK293T cells were seeded on 6-well plates (Corning) at 7.5×10^5 cells per well in DMEM plus GlutaMAX supplemented with 10% FBS. At 60% confluency 16 h after seeding, cells were transfected with 9 μL Lipofectamine RNAiMAX (Thermo Fisher Scientific) according to the manufacturer's protocol and 90 pmol ON-TARGETplus SMARTpool siRNAs (Horizon Discovery). One day after transfection, media was replaced with fresh DMEM plus GlutaMAX supplemented with 10% FBS. 2 days after transfection, cells were washed once with PBS and resuspended using TrypLE (Thermo Fisher Scientific) and DMEM plus GlutaMAX supplemented with 10% FBS. HEK293T cells were then seeded on 96-well plates (Corning) at 2.5×10^4 cells per well. Between 16 and 24 h after seeding, cells were

transfected at 60–80% confluency with 0.5 μL Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol and 200 ng prime editor plasmid, 66 ng pegRNA plasmid, 22 ng sgRNA plasmid (when required), and 5 pmol of the same ON-TARGETplus SMARTpool siRNAs used in the first transfection. For control conditions, cells were treated with non-targeting siRNAs in both transfections. For experiments in **FIG. 49F**, only the second transfection with PE components and siRNA was performed. Cells were cultured for 72 h after the second transfection before genomic DNA extraction. ***Real time quantitative PCR***

[778] To measure RNAi knockdown (**FIG. 49E**), RNA was isolated from HEK293T cells 72 h after the second siRNA transfection and converted to cDNA using the SYBR Green Fast Advanced Cells-to-CT Kit (Thermo Fisher Scientific) with cell lysis for 10-15 min using lysis solution containing 1:50 DNaseI to fully digest genomic DNA. All other steps were carried out according to the manufacturer's protocol. Each 20 μL qPCR reaction was performed in technical and biological triplicate with 500 nM of each primer, 2 μL cDNA, 1× SYBR Green (Thermo Fisher Scientific), and 10 μL Q5 High-Fidelity 2× Master Mix (New England BioLabs) on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories) with the following thermocycling conditions: 98 °C for 2 min, and 40 cycles of [98 °C for 15 s, 65 °C for 20 s, and 72 °C for 30 s]. β-actin (ACTB) served as a housekeeping gene to normalize the amount of cDNA in each qPCR reaction. Relative RNA abundances from gene knockdown were calculated in comparison to a non-targeting siRNA control by the $2^{-\Delta\Delta CT}$ method. A list of primers used for qPCR reactions is provided in Table 8.

Plasmid transfection dose titration in HEK293T cells

[779] For experiments in **FIG. 50B**, HEK293T cells were seeded on 96-well plates (Corning) at 1.6–1.8 × 10^4 cells per well in DMEM plus GlutaMAX supplemented with 10% FBS. Between 16 and 24 h after seeding, cells were transfected at 60–80% confluency with 0.5 μL Lipofectamine 2000 (Thermo Fisher Scientific) according to the manufacturer's protocol and 66 ng pegRNA plasmid, 0–200 ng PE2 plasmid, and 0–100 ng for MLH1dn or RFP plasmid. pUC19 filler plasmid was combined with PE2, MLH1dn, and RFP plasmids in different amounts to maintain a constant amount of total plasmid transfected (366 ng). In titrations varying the amount of total editor and *in trans* protein together, PE2 plasmid was used at a mass ratio of 2:1 with MLH1dn and RFP. Genomic DNA was isolated from cells 72 h after transfection.

Generation of* MLH1 *knock-out HeLa cell clones

[780] One clonal wild-type HeLa cell line and two clonal Δ*MLH1* HeLa lines were used to compare prime editing enhancement from MLH1dn expression versus *MLH1* knockout (**FIG. 50F**). To generate clonal lines, HeLa cells were seeded on 6-well plates (Corning) at 2.5×10^5 cells per well in DMEM plus GlutaMAX supplemented with 10% FBS. At 60% confluency 18 h after seeding, cells were transfected using 7.5 μL *Trans*IT-HeLa reagent (Mirus Bio) according to the manufacturer's protocol with 2 μg pLX_331-Cas9 (SpCas9 with blasticidin marker, Addgene #96924) and 500 ng sgRNA plasmid (spacer, 5′– GACAGTGGTGAACCGCATCG–3′ (SEQ ID NO: 367)). To make clonal wild-type HeLa cells as a control, 500 ng pUC19 plasmid was transfected instead of sgRNA plasmid. 24 h following transfection, 10 ng μL^{-1} blasticidin (Thermo Fisher Scientific) was added to each well to select for cells transfected with Cas9.

[781] Three days following transfection, cells were plated on 96-well plates at 1 cell per well with conditioned DMEM plus GlutaMAX supplemented with 10% FBS. Single clones were grown and expanded for 18 days. To verify that ΔMLH1 cells contain biallelic MLH1 frameshift mutations and that control cells contain the wild-type genotype, the MLH1 locus from clonal genomic DNA was sequenced on a MiSeq (Illumina) as described above. FASTQ sequencing files of MLH1 in HeLa clones are listed in Table 3. HeLa ΔMLH1 clone 1 contains MLH1 c.55_56insA and c.41_58delinsTAACTTCC alleles. HeLa ΔMLH1 clone 2 contains MLH1 c.55_56insA and c.20_66del alleles. All prime editing experiments with these clonal HeLa lines were performed as described above for HeLa cells.

Prime editing of contiguous substitutions and additional silent mutations

[782] Seven sets of prime edits that substitute 1–5 contiguous bases (35 edits total) were tested across five loci in HEK293T cells. Within each set of contiguous substitutions, all five edits altered at least one base within the seed region of the pegRNA protospacer (+1–3 nucleotides), at least one base within the PAM sequence of the pegRNA protospacer (+5 G or +6 G), or no bases within the seed region or PAM sequence at all. Because prime edits that alter the seed region or PAM sequence are more efficiently made, the design of these contiguous substitution edits controls for these confounding effects on editing efficiency, thereby enabling comparison of editing efficiency within each set.

[783] Six sets of prime edits that program a coding change with or without additional silent mutations (27 edits total) were tested across six gene targets in HEK293T cells. Each of the six

coding edits makes a transversion at one of the PAM nucleotides of the pegRNA protospacer (+5 G or +6 G). This design controls for confounding effects on editing efficiency (as explained above), allowing comparison of editing efficiency within each set. Silent mutations were designed to be close to (typically within 5-bp from) the intended coding edit to maximize interference of MMR recognition of the intended coding edit. The frequency of reads that contain the intended coding edit without indels and with or without any additional silent mutations was quantified using CRISPResso2 as described above.

Analysis of prime editor activity at Cas9 off-target sites

[784] Prime editor activity at known Cas9 off-target sites was determined by sequencing genomic DNA from HEK293T cells 3 days after transfection with plasmids encoding PE2, pegRNAs, and MLH1dn (when required) as described above. The top 4 off-target sites for each of the *HEK3*, *EMX1*, *FANCF*, and *HEK4* spacers previously detected by CIRCLE-seq (Tsai *et al.*, 2017) (16 sites total) were deep sequenced from genomic DNA samples as described above. To analyze off-target editing, reads were aligned to reference off-target amplicons using CRISPResso2 (Clement *et al.*, 2019) in standard mode with the parameters "-q 30" and "-w 10." Off-target reads were called as leniently as possible to capture all potential reverse transcription products. For each off-target reference amplicon, the nucleotide sequence 3′ of the Cas9 nick site (prime-editable target) was compared to the 3′ DNA flap sequence encoded by pegRNA reverse transcription. Counting from the 5′ ends, the minimum sequence of the 3′ DNA flap that deviates from prime-editable target sequence was designated as an off-target marker sequence. All reference-aligned reads that contain this off-target marker sequence directly 3′ of the Cas9 nick site (including indel-containing reads) were called as off-target reads. Off-target editing efficiencies were thus quantified as a percentage of (number of off-target reads)/(number of reference-aligned reads). For some amplicons, mismatch rates at the relevant editing position were comparable to rates at other positions in the amplicon, suggesting that context-specific sequencing errors may contribute to apparent off-target prime editing and therefore this conservative approach may overestimate the true rate of pegRNA-mediated editing at off-target sites.

Sequencing of microsatellite instability in genomic DNA

[785] Microsatellite instability was assessed in genomic DNA from HCT116 cells, monoclonal wild-type HAP1 cells, monoclonal HAP1 cells grown for 2 months (~60 cell division) following *MSH2* knockout, HeLa cells 3 days after transfection with plasmids encoding PE2–P2A–BSD, pegRNA, and MLH1dn when required. HeLa cell transfections were performed as described above. 17 mononucleotide repeats that are highly sensitive to MMR activity and are widely used to diagnosis MMR deficiency in tumors (Bacher *et al.*, 2004; Hempelmann *et al.*, 2015; Umar *et al.*, 2004) were deep sequenced from genomic DNA samples. The first PCR reaction (PCR1) amplified the microsatellite sequence of interest using primers (Integrated DNA Technologies) containing Illumina forward and reverse adapters. Each 20 µL PCR1 reaction was performed with 250 nM of each primer, 0.8 µL genomic DNA, 1× SYBR Green (Thermo Fisher Scientific), and 10 µL Q5 High-Fidelity 2× Master Mix (New England BioLabs) on a CFX96 Touch Real-Time PCR Detection System (Bio-Rad Laboratories) with the following thermocycling conditions: 98 °C for 3 min, 30 cycles of [98 °C for 15 s, 62 °C for 30 s, and 72 °C for 30 s], followed by 72 °C for 3 min. All 17 PCR1 products amplified from the same genomic DNA sample were pooled, purified with 0.8× AMPure XP beads (Beckman Coulter), and eluted in nuclease-free water. A list of primers used for PCR1 reactions is provided in Table 8. The subsequent PCR step (PCR2) added unique i7 and i5 Illumina barcode combinations to both ends of the PCR1 DNA amplicons to enable sample demultiplexing. Each 20 µL PCR2 reaction was performed with 500 nM of each barcoding primer, 25 ng of pooled PCR1 product, 1× SYBR Green, and 10 µL Q5 High-Fidelity 2× Master Mix on a CFX96 Touch Real-time PCR Detection System with the following thermocycling conditions: 98 °C for 2 min, 8 cycles of [98 °C for 15 s, 61 °C for 20 s, and 72 °C for 30 s], followed by 72 °C for 2 min. All PCR2 products were pooled, purified with 0.8× AMPure XP beads, and eluted in nuclease-free water. DNA amplicon libraries were quantified with a Qubit 3.0 Fluorometer (Thermo Fisher Scientific), then sequenced using the MiSeq Reagent Kit v2 (Illumina) with 300 single-read cycles. A list of FASTQ sequencing files generated in these experiments is provided in Table 7.

Quantification of microsatellite instability

[786] The seventeen microsatellites analyzed all consist of long homopolymers. To quantify the observed lengths of these microsatellites in a way that is robust against the high rate of

sequencing errors observed in homopolymers, each sequencing read was searched for the sequences expected to flank the homopolymers and then considered the final length of the homopolymer to be the distance between these flanking sequences. Specifically, for each locus, the longest homopolymer within the amplicon was identified, and 20-nt of the expected reference sequence on either side was recorded. Sequencing reads were demultiplexed into their loci of origin based on the first 20-nt of each read. Within reads for each locus, for each sequencing read, the first occurrences of sequences within Hamming distance 2 of the two flanking sequences were recorded. If both flanking sequences were located in the expected relative orientation within 50-nt of each, the distance between was recorded.

In vitro transcription of prime editor and MLH1dn mRNA used in iPSC and T cell experiments

[787] As described previously (Nelson *et al.*, 2021), plasmids were cloned to encode an inactivated T7 promoter followed by a 5′ untranslated region (UTR), Kozak sequence, coding sequences of PE2 or MLH1dn, and a 3′ UTR. T7 promoter inactivation prevents potential transcription from circular plasmid template during mRNA generation. These components together were PCR amplified with Phusion U Green Multiplex Master Mix (Thermo Fisher Scientific) using primers that correct T7 promoter inactivation and append a 119-nt poly(A) tail to the 3′ UTR. The resulting PCR product was purified with the QIAquick PCR Purification Kit (Thermo Fisher Scientific) and served as a template for subsequent *in vitro* transcription. PE2 and MLH1dn mRNAs were transcribed from these templates using the HiScribe T7 High-Yield RNA Synthesis Kit (New England BioLabs) with co-transcriptional capping by CleanCap AG (TriLink Biotechnologies) and full replacement of UTP with N^1-Methylpseudouridine-5′-triphosphate (TriLink Biotechnologies). Transcribed mRNAs were precipitated in 2.5 M lithium chloride (Thermo Fisher Scientific), washed twice in 70% ethanol, then dissolved in nuclease-free water. The resulting PE2 and MLH1dn mRNA was quantified with a NanoDrop One UV-Vis spectrophotometer (Thermo Fisher Scientific) and was stored at -80 °C.

Electroporation of human patient-derived induced pluripotent stem cells

[788] Prior to electroporation, 24-well culture plates (Thermo Fisher Scientific) were coated with 250 µL rhLaminin-521 (Thermo Fisher Scientific) diluted 1:40 in DPBS (Thermo Fisher Scientific) per well, and incubated at 37 °C in a 5% CO_2 incubator for 2 h. For electroporation, iPS cell colonies at 70–80% confluency were washed once with DPBS and dissociated in pre-

warmed Accutase (Innovative Cell Technologies) for 10 min at 37 °C in a 5% CO_2 incubator. Next, iPS cells were gently triturated, moved into a sterile 15 mL conical tube, then combined with an equal volume of DMEM/F12 (Thermo Fisher Scientific) to quench dissociation enzyme activity. Cells were pelleted at 300 *g* for 3 min and resuspended in StemFlex medium (Thermo Fisher Scientific) supplemented with 10 μM Y-27632 (Cayman Chemical). Cell counts and viability were determined using the Countess II FL Automated Cell Counter (Thermo Fisher Scientific). For electroporation using the NEON Transfection System 10 μL kit (Thermo Fisher Scientific), 2×10^5 iPS cells were pelleted at 300 *g* for 3 min and resuspended in 9 μL NEON Buffer R. The cell solution was combined with a 3 μL mixture of 1 μg PE2 mRNA, 90 pmol synthetic pegRNA (Integrated DNA Technologies), 60 pmol synthetic sgRNA (Synthego) when required, and 0–2 μg MLH1dn mRNA in NEON Buffer R. Synthetic pegRNAs and sgRNAs were dissolved in TE buffer (10 mM Tris-HCl, pH 8.0; 0.1 mM EDTA). Mock control electroporations were performed with 3 μL NEON Buffer R without any RNA added. Directly prior to electroporation, rhLaminin-521 was aspirated and immediately replaced with 250 μL pre-warmed StemFlex medium supplemented with 10 μM Y-27632 per 24-well. Next, 10 μL of the combined cell and RNA mixture was electroporated using the NEON Transfection System (Thermo Fisher Scientific) with the following parameters: 1400 V, 20 ms, one pulse. Cells were seeded immediately into rhLaminin-521-coated 24-well plates with 250 μL StemFlex medium supplemented with 10 μM Y-27632 per well. Media was changed the following day with 500 μL StemFlex medium supplemented with 5 μM Y-27632. 72 h following electroporation, media was changed to 500 μL StemFlex medium per well. Genomic DNA was extracted 96 h after electroporation by washing iPS cells once with DPBS, lysing with gDNA lysis buffer (10 mM Tris-HCl, pH 8.0; 0.05% SDS; 800 units μL^{-1} proteinase K (New England BioLabs)) at 37 °C for 2 h, followed by enzyme inactivation at 80 °C for 30 min. All iPSC electroporations were performed in technical duplicate and biological triplicate. Following amplicon sequencing of the edited *CDKL5* locus, frequencies of intended editing and indels were quantified with CRISPResso2 in HDR mode, as described above. Because patient-derived iPSCs were heterozygous for the c.1412delA allele, frequency of editable alleles with the intended edit was quantified as: (editing frequency – editing frequency in mock controls)/(100 – editing frequency in mock controls). Frequency of editable alleles with indels was quantified as described above: (total number of indel-containing reads)/(number of amplicon-aligned reads). The resulting

frequencies of editable alleles with the intended edit or indels were averaged between technical duplicates, and values from biological triplicates are shown.

Electroporation of primary human T cells

[789] Prior to electroporation, T cells were activated for 2 days with Dynabeads Human T-Expander CD3/CD28 (Thermo Fisher Scientific) and cultured at 37 °C and 5% CO_2 in T cell media (X-VIVO 15 Serum-free Hematopoietic Cell Medium (Lonza), supplemented with 5% AB human serum (Valley Biomedical), 1× GlutaMAX (Thermo Fisher Scientific), 12 mM N-acetyl-cysteine (Sigma Aldrich), 50 U mL^{-1} penicillin and 50 µg mL^{-1} streptomycin (Thermo Fisher Scientific), 300 IU mL^{-1} IL-2 (Peprotech), and 5 ng mL^{-1} recombinant human IL-7 (Peprotech) and IL-15 (Peprotech)). CD3/CD28 beads were removed from cells 5–7 h before electroporation. For electroporation using the NEON Transfection System 10 µL kit (Thermo Fisher Scientific), 3.0–3.5 × 10^5 cells per sample were pelleted by centrifugation for 5 min at 300 *g* and resuspended in 11 µL NEON Buffer T. The cell solution was added to a mix of 1 µg PE2 mRNA, 90 pmol synthetic pegRNA (Integrated DNA Technologies), 60 pmol synthetic sgRNA (Synthego), and 0–2 µg MLH1dn mRNA. Synthetic pegRNAs and sgRNAs were dissolved in TE buffer (10 mM Tris-HCl, pH 8.0; 0.1 mM EDTA). Mock control electroporations were performed with 3 µL NEON Buffer T without any RNA added. Electroporation on the NEON Transfection System (Thermo Fisher Scientific) was carried out using 10 µL NEON tips with the following parameters: 1,400 V, 10 ms, three pulses. Cells were plated in 600 µL fresh T cell media in a 24-well plate. 2 days after electroporation, cell counts and viability were determined using the Countess II Automated Cell Counter (Thermo Fisher Scientific) and 1 mL fresh T cell media was added to cells. 4 days after electroporation, cells were pelleted by centrifugation for 5 min at 300 *g* and genomic DNA was isolated using the PureLink Genomic DNA Mini Kit (Thermo Fisher Scientific) following the "mammalian cells lysate" protocol with elution in nuclease-free water.

Quantification And Statistical Analysis

[790] The number of independent biological replicates and technical replicates for each experiment are described in the Brief Description of the Drawings or the Methods details section. In FIG. 44G, a non-parametric Mann-Whitney *U* test was used to compare prime editing data in HEK293T cells with prime editing data in HeLa, K562, and U2OS cells.

This Example refers to Tables 5-7, which are provided below:

Table 5

Spacer_Sequence_Sense +5'G	Forward_CRISPRi_Library_Oligo	Reverse_CRISPRi_Library_Oligo	Counts_In(oligo-pool	CRISPRi_guide_ID_1	Gene_Symbol_1	Transcript_1	CRISPRi_guide_ID_2	Gene_Symbol_2	Transcript_2	CRISPRi_guide_ID_3	Gene_Symbol_3	Transcript_3	CRISPRi_guide_ID_4	Gene_Symbol_4	Transcript_4
GCCCACCTCAAGGCAAGA AAG (SEQ ID NO: 463)	TATGAACCACTAAGGCGTccaccgtgtgGCCCACCTCAAGGC (AACAAAGgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2038)	TGACCAGGCAGAGTTTACGCAGCTTAGCTCTTAAACCTTTGTTG CCTGAGGTGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3699)	1	ACTL6A_+_179280744.23-P1P2	ACTL6A	P1P2	na	na	na	na	na	na	na	na	na
GCGGGCGCGACTCCTACC CTA (SEQ ID NO: 464)	TATGAACCACTAAGGCGTccaccgtgtgGCGGGCGCGACTCC (TACCCTAgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2037)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGGGTA (GGAGTCGCGCGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3610)	1	ACTL6A_+_179280840.23-P1P2	ACTL6A	P1P2	na	na	na	na	na	na	na	na	na
GGACGAGTTAGCCCTTTAG GGT (SEQ ID NO: 465)	TATGAACCACTAAGGCGTccaccgtgtgCGCCAGTTTAGCCC (TTAGGGTgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2038)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCTAAG (GGCTAACTGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3611)	1	ACTL6A_-_179280860.23-P1P2	ACTL6A	P1P2	na	na	na	na	na	na	na	na	na
GACTGGGTTCCGGTTGCAG CAC (SEQ ID NO: 466)	TATGAACCACTAAGGCGTccaccgtgtgGACTGGGTTCCGGT (ECGGGCACgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2039)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTGCCGC (AGCCGGACCCAGTGCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3612)	1	ACTR5_+_37377152.23-P1P2	ACTR5	P1P2	na	na	na	na	na	na	na	na	na
GGAACGCGCGCTCCAAGA TGG (SEQ ID NO: 467)	TATGAACCACTAAGGCGTccaccgtgtgGGAACGCGCGCTCC (AAGATGgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2040)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATCTTG (GAGCGCGCGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3613)	1	ACTR5_-_37377127.23-P1P2	ACTR5	P1P2	na	na	na	na	na	na	na	na	na
GGGCACACGCGCTCCGCG GAAC (SEQ ID NO: 468)	TATGAACCACTAAGGCGTccaccgtgtgCGGCACACGCGCTCC (GAACgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2041)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGGGT (GACGCCCACCGACTCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3614)	1	ACTR5_+_37377138.23-P1P2	ACTR5	P1P2	na	na	na	na	na	na	na	na	na
GAGTCGTGGCTGCAGCG CTG (SEQ ID NO: 469)	TATGAACCACTAAGGCGTccaccgtgtgGAGTCGTGGCTGC (AGCGCTgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2042)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGACGAGCT (CGAGCCCACCGACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3615)	1	ACTR8_+_53916157.23-P1P2	ACTR8	P1P2	na	na	na	na	na	na	na	na	na
GGCTGAGGTCGAGAAGGTT GGT (SEQ ID NO: 470)	TATGAACCACTAAGGCGTccaccgtgtgGGCTGAGGTCGAGA (AGGTGGTgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2043)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCAACCT (CTCGCCTCAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3616)	1	ACTR8_+_53916142.23-P1P2	ACTR8	P1P2	na	na	na	na	na	na	na	na	na
GTCTCCGGGCATAATGAC TCC (SEQ ID NO: 471)	TATGAACCACTAAGGCGTccaccgtgtgGTCTCCGGGCATAA (TGACCgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2044)	TGACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGCCTAT (ATGGCCGGAGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3617)	1	ACTR8_+_53916118.23-P1P2	ACTR8	P1P2	na	na	na	na	na	na	na	na	na
GCCACTATGGACAGGTAA AG (SEQ ID NO: 472)	TATGAACCACTAAGGCGTccaccgtgtgGCCACTATGGACA (CGTAAAGgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2045)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTTACCT (CTCCATAGTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3618)	1	AIODA_+_85346.23-P1P2	AIODA	P1P2	na	na	na	na	na	na	na	na	na
GGACAGTCTTGGACAGCCA CTA (SEQ ID NO: 473)	TATGAACCACTAAGGCGTccaccgtgtgGGACAGTCTTGGAC (ACCAGCTAgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2046)	TGACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGTGGT (GTCCAGAGTGTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3619)	1	AIODA_+_85396.23-P1P2	AIODA	P1P2	na	na	na	na	na	na	na	na	na
GTTCTCTGGCGTGAGACTTG GTG (SEQ ID NO: 474)	TATGAACCACTAAGGCGTccaccgtgtgGTTCTGGCGTGAG (ACTTGCAgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2047)	TCAGCCAGAGACTTTACGCAGCTTAGCTCTTAAACGTGCAAGTC (TCAGGCCAGAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3620)	1	AIODA_+_85390.23-P1P2	AIODA	P1P2	na	na	na	na	na	na	na	na	na
GAATCCCGGCAGTCACGCT GTG (SEQ ID NO: 475)	TATGAACCACTAAGGCGTccaccgtgtgGAATCCCGGCAGTCA (CCCTGGgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2048)	TGACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCAGGGGT (GACTGCGGGAGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3621)	1	ALK3H2_-_109537258.23-P1P2	ALK3H2	P1P2	na	na	na	na	na	na	na	na	na
GAGGGCACCACAGGGTGAC TGC (SEQ ID NO: 476)	TATGAACCACTAAGGCGTccaccgtgtgGAGGGCACCACAGGG (TGACTGGgttaagagttaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2049)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGCAGTCA (CCCTGGTGTGGCGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3622)	1	ALK3H2_+_109312418.23-P1P2	ALK3H2	P1P2	na	na	na	na	na	na	na	na	na

GTCCCACCACGAGGCCACCA (SEQ ID NO: 477)	TATGAACCACTAAGGCGTccaccttgttgGTCCCACCACGAGGCCACCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2050)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGTGGCCTCGTGGTGGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3623)	1	ALKBH2_+_109531250.23-P1P2	ALKBH2	P1P2	na	na	na	na	na	na	na	na	na
GAGACTGCGGAGTGGGTCAG (SEQ ID NO: 478)	TATGAACCACTAAGGCGTccaccttgttgGAGACTGCGGAGTGGGTCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2051)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGACCCACTCCGCAGTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3624)	1	ALKBH3_-_43902421.23-P1P2	ALKBH3	P1P2	na	na	na	na	na	na	na	na	na
GAGGTTCCGATCACAGACTG (SEQ ID NO: 479)	TATGAACCACTAAGGCGTccaccttgttgGAGGTTCCGATCACAGACTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2052)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGTCTGTGATCGGAACCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3625)	1	ALKBH3_-_43902408.23-P1P2	ALKBH3	P1P2	na	na	na	na	na	na	na	na	na
GGCAAACCCGGTAGGACTTG (SEQ ID NO: 480)	TATGAACCACTAAGGCGTccaccttgttgGGCAAACCCGGTAGGACTTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2053)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAAGTCCTACCGGGTTTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3626)	1	ALKBH3_+_43902434.23-P1P2	ALKBH3	P1P2	na	na	na	na	na	na	na	na	na
GCGGCGGGCCCATCCGGAAC (SEQ ID NO: 481)	TATGAACCACTAAGGCGTccaccttgttgGCGGCGGGCCCATCCGGAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2054)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTCCGGATGGGCCCGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3627)	1	ALYREF_+_79849261.23-P1P2	ALYREF	P1P2	na	na	na	na	na	na	na	na	na
GCGGGCGCGGAATCGGGCAT (SEQ ID NO: 482)	TATGAACCACTAAGGCGTccaccttgttgGCGGGCGCGGAATCGGGCATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2055)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATGCCCGATTCCGCGCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3628)	1	ALYREF_-_79849458.23-P1P2	ALYREF	P1P2	na	na	na	na	na	na	na	na	na
GGGAGGCGGCAGGAACCGAC (SEQ ID NO: 483)	TATGAACCACTAAGGCGTccaccttgttgGggAGGCGGCAGGAACCGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2056)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCGGTTCCTGCCGCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3629)	1	ALYREF_+_79849212.23-P1P2	ALYREF	P1P2	na	na	na	na	na	na	na	na	na
GCGTATCTGCACCGGCCCGA (SEQ ID NO: 484)	TATGAACCACTAAGGCGTccaccttgttgGCGTATCTGCACCGGCCCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2057)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGGGCCGGTGCAGATACGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3630)	1	APEX1_+_20923418.23-P1P2	APEX1	P1P2	na	na	na	na	na	na	na	na	na
GGAGGAGCTAGGCTGCCATC (SEQ ID NO: 485)	TATGAACCACTAAGGCGTccaccttgttgGGAGGAGCTAGGCTGCCATCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2058)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGATGGCAGCCTAGCTCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3631)	1	APEX1_-_20923425.23-P1P2	APEX1	P1P2	na	na	na	na	na	na	na	na	na
GTCTTTGTGCTCGGGTTAGG (SEQ ID NO: 486)	TATGAACCACTAAGGCGTccaccttgttgGTCTTTGTGCTCGGGTTAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2059)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTAACCCGAGCACAAAGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3632)	1	APEX1_-_20923407.23-P1P2	APEX1	P1P2	na	na	na	na	na	na	na	na	na
GAGTTCGCTCGCGCCTAGGT (SEQ ID NO: 487)	TATGAACCACTAAGGCGTccaccttgttgGAGTTCGCTCGCGCCTAGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2060)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCTAGGCGCGAGCGAACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3633)	1	APEX2_-_55026823.23-P1P2	APEX2	P1P2	na	na	na	na	na	na	na	na	na
GGCTCGCGCCTAGGTTGGCG (SEQ ID NO: 488)	TATGAACCACTAAGGCGTccaccttgttgGGCTCGCGCCTAGGTTGGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2061)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCAACCTAGGCGCGAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3634)	1	APEX2_-_55026828.23-P1P2	APEX2	P1P2	na	na	na	na	na	na	na	na	na
GTCCCAGCCCGCGCCAACCT (SEQ ID NO: 489)	TATGAACCACTAAGGCGTccaccttgttgGTCCCAGCCCGCGCCAACCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2062)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGGTTGGCGCGGGCTGGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3635)	1	APEX2_+_55026814.23-P1P2	APEX2	P1P2	na	na	na	na	na	na	na	na	na
GCACCGAGGGAGGGCCTCAA (SEQ ID NO: 490)	TATGAACCACTAAGGCGTccaccttgttgGCACCGAGGGAGGGCCTCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2063)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGAGGCCCTCCCTCGGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3636)	1	APLF_+_68694742.23-P1P2	APLF	P1P2	na	na	na	na	na	na	na	na	na
GCGCTGCTGGGAGTAAGTGT (SEQ ID NO: 491)	TATGAACCACTAAGGCGTccaccttgttgGCGCTGCTGGGAGTAAGTGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2064)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACACTTACTCCCAGCAGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3637)	1	APLF_-_68694969.23-P1P2	APLF	P1P2	na	na	na	na	na	na	na	na	na
GGGCCGGGCTCTGAGAGGAC (SEQ ID NO: 492)	TATGAACCACTAAGGCGTccaccttgttgGGGCCGGGCTCTGAGAGGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2065)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCCTCTCAGAGCCCGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3638)	1	APLF_-_68694726.23-P1P2	APLF	P1P2	na	na	na	na	na	na	na	na	na
GAAAGGTAAAATGGATCCGG (SEQ ID NO: 493)	TATGAACCACTAAGGCGTccaccttgttgGAAAGGTAAAATGGATCCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2066)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGATCCATTTTACCTTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3639)	1	APOBEC1_+_7818434.23-P1P2	APOBEC1	P1P2	na	na	na	na	na	na	na	na	na

GCAACTTATGACTCCAGA GG (SEQ ID NO: 484)	TATGAAGCACTAAG3GCGTccaactgtggCAACTTATGACTCC AGTCATAAGTTGCCAACAAGGTGGACGCCTTAGTGGTCATA (SEQ ID NO: 3640)	TCAGCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACCCTCTGG _7816895 23-P1P2	APOBEC1_+ 7816895 23-P1P2	APOBEC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACACAACTTATGACTCC AG (SEQ ID NO: 485)	TATGAACCACTAAGGCGTccaactgtggCACAACTTATGA (SEQ ID NO: 3641)	TCAGCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACCTGGAGT CATAAGTTGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3641)	APOBEC1_+ 7816821 23-P1P2	APOBEC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCTTGACAATTCTCAAAG GG (SEQ ID NO: 486)	TATGAACCACTAAGCGCGTccaactgtgACTTGTCACAATCTC (SEQ ID NO: 3642)	TCAACAGCAGACTTTAGCGCAACAAGGTTGACTCTTAAACCCCTTTGA GATTTGTCACAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3642)	APOBEC2_+ 4102119 42 3-P1P2	APOBEC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCATTGGCCCACAAGC AAG (SEQ ID NO: 487)	TATGAAGCACTAAGGCGTccaactgtgCTCCATTGGCCCAG AAGGGAAGgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3643)	TCAGCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACCTTCCTTC TGGGCCATGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3643)	APOBEC2_- 4102119 42 3-P1P2	APOBEC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGTGGCTCAAGGAGTCA CTG (SEQ ID NO: 488)	TATGAACCACTAAGGCGTccaactgtgGTGGCTCAAGGA GTCACTGgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3644)	TCAGCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACCAGTGACT CCTGAGCCACAGCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3644)	APOBEC2_+ 4102106 5 2 3-P1P2	APOBEC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACAGACCAAGGAAGCGA GAA (SEQ ID NO: 489)	TATGAACCACTAAGGCGTccaactgtgGACAGACCAAGGAA CCGAAGAAgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3645)	TCAGCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACTTCTCGGT TCCTGGTCTGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3645)	APOBEC3A_ 39353687 2 3-P1P2	APOBEC3A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGCAAGTCGCCAAGAGC GGG (SEQ ID NO: 490)	TATGAAGCACTAAGCGCGTccaactgtgGAGCAAGTCGCCAA TCGCGAGTgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3646)	TCAGCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACTCCCGCTCT TCGCGACTTGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3646)	APOBEC3A_ 39353693 2 3-P1P2	APOBEC3A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCTCCTTGCGACTTGCT CA (SEQ ID NO: 491)	TATGAACCACTAAGGCGTccaactgtgGCGCTCTTGCGACC ITCGTCAgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3647)	TCACCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACTGAGGCAA GTCGCAAGACGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3647)	APOBEC3A_ 39353687 23-P1P2	APOBEC3A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGCCTGAGTGGGATTC GTG (SEQ ID NO: 492)	TATGAACCACTAAGGCGTccaactgtgGAGCCTGAGTGGG ATTCGTGgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3648)	TCAGCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACCACGAATC CCAGTCAAGGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3648)	APOBEC3B_ + 39378603 23-P1P2	APOBEC3B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGACAAAGACGCCTG ACT (SEQ ID NO: 493)	TATGAAGCACTAAGGCGTccaactgtgCAGGACAAAGACG CCTGACTgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3649)	TCAGCAGCAGACTTTAGCGCAGGTTAGCTCTTAAACAGTCAGGG CTCTTTGTCCTGCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3649)	APOBEC3B_ + 39378613 23-P1P2	APOBEC3B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTCAAAAAAAGAGCGG GAC (SEQ ID NO: 494)	TATGAACCACTAAGGCGTccaactgtgGTTCAAAAAAAGAG CGGGGACgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3650)	TCACCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACGGTCCCGC TCTTTTTTTGAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3650)	APOBEC3B_ 39379432 2 3-P1P2	APOBEC3B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGAGCTCCCCAGGAAT AGT (SEQ ID NO: 495)	TATGAACCACTAAGGCGTccaactgtgGCAGAGCTCCCAG GAATAGTgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3651)	TCACCAGCAGACTTTAGCGCAGGTTAGCTCTTAAACACTATTCCC TGGGAAGCTCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3651)	APOBEC3C_ + 39410687 23-P1P2	APOBEC3C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGCGCTTGAGAAAAAG AGT (SEQ ID NO: 496)	TTAAGAACACTAAGGCGTccaactgtgCGAGCGCTTGAGA AAAGAGTgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3652)	TCAGCAGCAGACTTTAGCGCAGGTTAGCTCTTAAACGATCTTTT CTTGAAGCGCTCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3652)	APOBEC3C_ 39410338 2 3-P1P2	APOBEC3C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTCAGAAAAAGAGTGGG ACA (SEQ ID NO: 497)	TATGAACCACTAAGGCGTccaactgtgGTTCAGAAAAAGAGT GGGGACAgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3653)	TCACCAGCAGACTTTAGCGCAGGTTAGCTCTTAAACGTGTCCCAG TCTTTTCTGAAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3653)	APOBEC3C_ 39410342 2 3-P1P2	APOBEC3C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAAGAGCAGGAAAAAGA GCC (SEQ ID NO: 498)	TTAAGAACACTAAGGCGTccaactgtgCAAGAGCAGGAAA AAGAGCCgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3654)	TCACCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACGGCTTTT TTCCTGCTCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3654)	APOBEC3F_ + 39437128 23-P1P2	APOBEC3F	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACCTGCCATGACTTTC GTG (SEQ ID NO: 499)	TATGAAGCACTAAGGCGTccaactgtgGGCCTGCCATGA CTTTCGTgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3655)	TCAGCAGCAGACTTTAGCGCAGGTTAGCTCTTAAACCCGAAAG TCAGTGGCAGTGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3655)	APOBEC3F_ + 39437228 23-P1P2	APOBEC3F	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGGCGTGTCTCTGAA ACC (SEQ ID NO: 500)	TATGAACCACTAAGGCGTccaactgtgGGAAGGCGTGTCT TGAAACCgttaagtgcttaagtgcGTAAAGTCTGCTGGTGA (SEQ ID NO: 3656)	TCACCAGCAGACTTTAGCGCAAGGTTAGCTCTTAAACGGTTTCAG GACAGGCCTCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3656)	APOBEC3F_ + 39436895 2 3-P1P2	APOBEC3F	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GACCTGGGCTGGGACAGATG (SEQ ID NO: 511)	TATGAACCACTAAGGCGTccaccttgttgGACCTGGGCTGGGACAGATGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2084)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCATCTGTCCCAGCCCAGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3657)	1	APOBEC3G_+_39473587.23-P1P2	APOBEC3G	P1P2	na	na	na	na	na	na	na	na	na
GAGGAGCCTGGCTGGGAGTA (SEQ ID NO: 512)	TATGAACCACTAAGGCGTccaccttgttgGAGGAGCCTGGCTGGGAGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2085)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACTCCCAGCCAGGCTCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3658)	1	APOBEC3G_+_39473517.23-P1P2	APOBEC3G	P1P2	na	na	na	na	na	na	na	na	na
GTCCAGAGTTCAGGAATGGC (SEQ ID NO: 513)	TATGAACCACTAAGGCGTccaccttgttgGTCCAGAGTTCAGGAATGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2086)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCATTCCTGAACTCTGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3659)	1	APOBEC3G_+_39473688.23-P1P2	APOBEC3G	P1P2	na	na	na	na	na	na	na	na	na
GACTTGCCCCTGCTTGTCTG (SEQ ID NO: 514)	TATGAACCACTAAGGCGTccaccttgttgGACTTGCCCCTGCTTGTCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2087)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGACAAGCAGGGGCAAGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3660)	1	APOBEC3H_+_39493291.23-P1P2	APOBEC3H	P1P2	na	na	na	na	na	na	na	na	na
GCAGGGCAAAGAGCCTGGGT (SEQ ID NO: 515)	TATGAACCACTAAGGCGTccaccttgttgGCAGGGCAAAGAGCCTGGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2088)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCCAGGCTCTTTGCCCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3661)	1	APOBEC3H_+_39493488.23-P1P2	APOBEC3H	P1P2	na	na	na	na	na	na	na	na	na
GTAGTGCCTCAGACAAGCAG (SEQ ID NO: 516)	TATGAACCACTAAGGCGTccaccttgttgGTAGTGCCTCAGACAAGCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2089)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGCTTGTCTGAGGCACTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3662)	1	APOBEC3H_-_39493307.23-P1P2	APOBEC3H	P1P2	na	na	na	na	na	na	na	na	na
GCAGCAAGGGCTTTCTAACT (SEQ ID NO: 517)	TATGAACCACTAAGGCGTccaccttgttgGCAGCAAGGGCTTTCTAACTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2090)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGTTAGAAAGCCCTTGCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3663)	1	APOBEC4_+_183622401.23-P1P2	APOBEC4	P1P2	na	na	na	na	na	na	na	na	na
GGAAAGCCCTTGCTGAGGTG (SEQ ID NO: 518)	TATGAACCACTAAGGCGTccaccttgttgGGAAAGCCCTTGCTGAGGTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2091)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCACCTCAGCAAGGGCTTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3664)	1	APOBEC4_-_183622430.23-P1P2	APOBEC4	P1P2	na	na	na	na	na	na	na	na	na
GGGTAGCCACACCTCAGCAA (SEQ ID NO: 519)	TATGAACCACTAAGGCGTccaccttgttgGGGTAGCCACACCTCAGCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2092)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGCTGAGGTGTGGCTACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3665)	1	APOBEC4_+_183622414.23-P1P2	APOBEC4	P1P2	na	na	na	na	na	na	na	na	na
GCAGCCCAAGGTAGTAACAG (SEQ ID NO: 520)	TATGAACCACTAAGGCGTccaccttgttgGCAGCCCAAGGTAGTAACAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2093)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGTTACTACCTTGGGCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3666)	1	APTX_-_33001417.23-P1	APTX	P1	na	na	na	na	na	na	na	na	na
GCCCAGCCAGCAGAAGAGAT (SEQ ID NO: 521)	TATGAACCACTAAGGCGTccaccttgttgGCCCAGCCAGCAGAAGAGATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2094)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATCTCTTCTGCTGGCTGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3667)	1	APTX_-_33001547.23-P1	APTX	P1	na	na	na	na	na	na	na	na	na
GGTCCGTCTCCGACTTCTGG (SEQ ID NO: 522)	TATGAACCACTAAGGCGTccaccttgttgGGTCCGTCTCCGACTTCTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2095)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCAGAAGTCGGAGACGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3668)	1	APTX_+_33001563.23-P1	APTX	P1	na	na	na	na	na	na	na	na	na
GCACTCTCAGGCTCGCTGAG (SEQ ID NO: 523)	TATGAACCACTAAGGCGTccaccttgttgGCACTCTCAGGCTCGCTGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2096)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCAGCGAGCCTGAGAGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3669)	1	APTX_-_33025057.23-P2	APTX	P2	na	na	na	na	na	na	na	na	na
GCGGCGCTCGTGGAATCTTC (SEQ ID NO: 524)	TATGAACCACTAAGGCGTccaccttgttgGCGGCGCTCGTGGAATCTTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2097)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAAGATTCCACGAGCGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3670)	1	APTX_+_33025069.23-P2	APTX	P2	na	na	na	na	na	na	na	na	na
GCTGCGCCGAAAGGCCTTGT (SEQ ID NO: 525)	TATGAACCACTAAGGCGTccaccttgttgGCTGCGCCGAAAGGCCTTGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2098)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACAAGGCCTTTCGGCGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3671)	1	APTX_+_33025113.23-P2	APTX	P2	na	na	na	na	na	na	na	na	na
GCCGGCCCGGAGCGTTTGTG (SEQ ID NO: 526)	TATGAACCACTAAGGCGTccaccttgttgGCCGGCCCGGAGCGTTTGTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2099)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCACAAACGCTCCGGGCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3672)	2	AQR_+_35261919.23-P1P2	AQR	P1P2	na	na	na	na	na	na	na	na	na
GCTGTTGTACGGGTTCCCTA (SEQ ID NO: 527)	TATGAACCACTAAGGCGTccaccttgttgGCTGTTGTACGGGTTCCCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2100)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGGGAACCCGTACAACAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3673)	2	AQR_+_35261962.23-P1P2	AQR	P1P2	na	na	na	na	na	na	na	na	na

GGGCGCCCGGCAGTAAC TTA (SEQ ID NO: 529)	TATGAACCACTAAGGCGTcaacctgtgGGGCCCGGGCAG TAACTTAgttaagtgctaagcgtCGTAAAGTTCTGCTGGTCA (SEQ ID NO: 2101)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTAAGTTAC TGCCGGCGCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3674)	2	PI1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAAGCGGGAGAGTCACA GCG (SEQ ID NO: 528)	TATGAACCACTAAGGCGTcaacctgtgGAAAGCGGGAGAGT CACAGCGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2102)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACGGCTGTG ACTGTCGCGGTTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3675)	1	PI1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCTTCGCCTGGGAATG AGG (SEQ ID NO: 526)	TATGAACCACTAAGGCGTcaacctgtgCCCTTGCCTGGG AGTGAGGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2103)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCTCATTC CCAGGCAAGGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3676)	1	PI1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTTCTCCGGCAGCAGAA AG (SEQ ID NO: 531)	TATGAACCACTAAGGCGTcaacctgtgGTTTCTCCGGCAG CAGAAAGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2104)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCTTTCGTGC TGCGGCGAGAAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3677)	1	PI1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACCGGGGACCATGCT GGG (SEQ ID NO: 532)	TATGAACCACTAAGGCGTcaacctgtgCACCGGGGACCATG CTGGGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2105)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCCAGCAT CTGGTCCCGGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3678)	1	PI1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGAGGCGCGCTCTCAA GGA (SEQ ID NO: 533)	TATGAACCACTAAGGCGTcaacctgtgGCCGAGGCGCGCT CTCAAGGAgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2106)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTCCTTGAG ATGCGCGCTCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3679)	1	PI1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGCGTGCCATGATCG CCG (SEQ ID NO: 534)	TATGAACCACTAAGGCGTcaacctgtgCGCGCGTGCGGCCATG ATgcgcgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2107)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCTCGCGAT CATGGCGCGCGCGGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3680)	1	PI1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGAACCACGCACCCGT CGG (SEQ ID NO: 535)	TATGAACCACTAAGGCGTcaacctgtgGCAGAACCACGCA CCCGTCGGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2108)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCGCAGGG CtTGCTCGGTCTTGCGCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3681)	1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCATGATGCGAGAGCTA CGG (SEQ ID NO: 536)	TATGAACCACTAAGGCGTcaacctgtgGCCATGATGCGGGA TGCGCATCATGGCGCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 2109)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCGTATGC TGCGCATCATGGCGCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3682)	1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGTTGAGCCAGCGTCAT TGA (SEQ ID NO: 537)	TATGAAGCACTAAGGCGTcaacctgtgGGAGGTTGAGCCA GCGTCATGAgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2110)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTCAATCAG CTGCTCAGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3683)	1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACCGGGGGGGGGCAGC GCCG (SEQ ID NO: 538)	TATGAACCACTAAGGCGTcaacctgtgGAACCGGGGGGGGC AGCGCCGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2111)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCGGGGGCT GCCCCGCCGGTTGCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3684)	1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCGCGCCGCCGCAG GAAT (SEQ ID NO: 539)	TATGAACCACTAAGGCGTcaacctgtgGCCGCCGCCGCC GCAGGAATgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2112)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACATTCCTCG GCCGGCGCGGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3685)	1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCAGTCATGCGCGCGG CGG (SEQ ID NO: 540)	TATGAAGCACTAAGGCGTcaacctgtgGCTCAGTCATGCGG CGGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2113)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCGCGCGG CCCCATGACTGAGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3686)	1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGAGCGCAGACAGCGG GCG (SEQ ID NO: 541)	TATGAACCACTAAGGCGTcaacctgtgGAAGAGCGCAGACA GAGCGCGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2114)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACGCGCGCGG CGTGTCGCGTCTTCGCAAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3687)	1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACGCAAAGACCGCACAG ACC (SEQ ID NO: 542)	TATGAACCACTAAGGCGTcaacctgtgGAACGCAAAGACCG CACAGACCgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2115)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTGTCTTGT GCGTCGTTTGCGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3688)	1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGCACAGACCCGGCGA GGA (SEQ ID NO: 543)	TATGAACCACTAAGGCGTcaacctgtgGCTCGCACAGACCC GGAGGTGTgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2116)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTCCTCGGC GCTCGCACAGACCCGGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3689)	1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGCGCAGAGCTTCGTG CAG (SEQ ID NO: 544)	TATGAACCACTAAGGCGTcaacctgtgGCCGGCGCAGAGCT TCGTGCAGgttaagtgctaagcgtaagcgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2117)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACGTCGCACG ACGTGTCGCGGCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3690)	1	ASF1A	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGGCTGTAGTTTAGTGAAAT (SEQ ID NO: 545)	TATGAACCACTAAGGCGTccaccttgttgGGGCTGTAGTTTAGTGAAATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2118)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATTTCACTAAACTACAGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3691)	1	ASF1A_-_119215390.23-P1P2	ASF1A	P1P2	na	na	na	na	na	na	na	na	na
GGTTCAGGTGAACAATGTAG (SEQ ID NO: 546)	TATGAACCACTAAGGCGTccaccttgttgGGTTCAGGTGAACAATGTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2119)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTACATTGTTCACCTGAACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3692)	1	ASF1A_-_119215607.23-P1P2	ASF1A	P1P2	na	na	na	na	na	na	na	na	na
GAAGAAGCCGTGGCCGCGGG (SEQ ID NO: 547)	TATGAACCACTAAGGCGTccaccttgttgGAAGAAGCCGTGGCCGCGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2120)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCGCGGCCACGGCTTCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3693)	1	ATM_-_108093839.23-P1P2	ATM	P1P2	na	na	na	na	na	na	na	na	na
GTAGGTAGCTGCGTGGCTAA (SEQ ID NO: 548)	TATGAACCACTAAGGCGTccaccttgttgGTAGGTAGCTGCGTGGCTAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2121)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTAGCCACGCAGCTACCTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3694)	1	ATM_-_108093813.23-P1P2	ATM	P1P2	na	na	na	na	na	na	na	na	na
GTTGAACCGGAAGCGGGAGT (SEQ ID NO: 549)	TATGAACCACTAAGGCGTccaccttgttgGTTGAACCGGAAGCGGGAGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2122)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACTCCCGCTTCCGGTTCAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3695)	1	ATM_-_108093795.23-P1P2	ATM	P1P2	na	na	na	na	na	na	na	na	na
GGGAGGGACTGCGGCAGTGT (SEQ ID NO: 550)	TATGAACCACTAAGGCGTccaccttgttgGGGAGGGACTGCGGCAGTGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2123)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACACTGCCGCAGTCCCTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3696)	1	ATP13A2_+_17338415.23-P1P2	ATP13A2	P1P2	na	na	na	na	na	na	na	na	na
GGGCTGGAGCAGGGCTGACG (SEQ ID NO: 551)	TATGAACCACTAAGGCGTccaccttgttgGGGCTGGAGCAGGGCTGACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2124)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTCAGCCCTGCTCCAGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3697)	1	ATP13A2_-_17338380.23-P1P2	ATP13A2	P1P2	na	na	na	na	na	na	na	na	na
GTGGGCCACCAGGCTCGGCG (SEQ ID NO: 552)	TATGAACCACTAAGGCGTccaccttgttgGTGGGCCACCAGGCTCGGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2125)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCGAGCCTGGTGGCCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3698)	1	ATP13A2_-_17338412.23-P1P2	ATP13A2	P1P2	na	na	na	na	na	na	na	na	na
GAGCCCAAGGCGGTGAGTAT (SEQ ID NO: 553)	TATGAACCACTAAGGCGTccaccttgttgGAGCCCAAGGCGGTGAGTATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2126)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATACTCACCGCCTTGGGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3699)	1	ATP13A5_+_193096393.23-P1P2	ATP13A5	P1P2	na	na	na	na	na	na	na	na	na
GCTCGCCGGTTGAGTTCAGA (SEQ ID NO: 554)	TATGAACCACTAAGGCGTccaccttgttgGCTCGCCGGTTGAGTTCAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2127)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTGAACTCAACCGGCGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3700)	1	ATP13A5_+_193096510.23-P1P2	ATP13A5	P1P2	na	na	na	na	na	na	na	na	na
GGCTTTGCTCAACCAGGGAG (SEQ ID NO: 555)	TATGAACCACTAAGGCGTccaccttgttgGGCTTTGCTCAACCAGGGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2128)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCCTGGTTGAGCAAAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3701)	1	ATP13A5_+_193096462.23-P1P2	ATP13A5	P1P2	na	na	na	na	na	na	na	na	na
GAGACGCCGGGAACCCGCGT (SEQ ID NO: 556)	TATGAACCACTAAGGCGTccaccttgttgGAGACGCCGGGAACCCGCGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2129)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGCGGGTTCCCGGCGTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3702)	2	ATR_+_142297576.23-P1P2	ATR	P1P2	na	na	na	na	na	na	na	na	na
GCCGGGAACCCGCGTTGGCG (SEQ ID NO: 557)	TATGAACCACTAAGGCGTccaccttgttgGCCGGGAACCCGCGTTGGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2130)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCAACGCGGGTTCCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3703)	2	ATR_+_142297571.23-P1P2	ATR	P1P2	na	na	na	na	na	na	na	na	na
GGCCCGACGGAGCCGTGTGG (SEQ ID NO: 558)	TATGAACCACTAAGGCGTccaccttgttgGGCCCGACGGAGCCGTGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2131)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACACGGCTCCGTCGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3704)	2	ATR_-_142297649.23-P1P2	ATR	P1P2	na	na	na	na	na	na	na	na	na
GCACTCGCGGCGGAGGCAAG (SEQ ID NO: 559)	TATGAACCACTAAGGCGTccaccttgttgGCACTCGCGGCGGAGGCAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2132)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTGCCTCCGCCGCGAGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3705)	1	ATRIP_-_48488160.23-P1P2	ATRIP	P1P2	na	na	na	na	na	na	na	na	na
GTCGGCGCTGTCGGATACTT (SEQ ID NO: 560)	TATGAACCACTAAGGCGTccaccttgttgGTCGGCGCTGTCGGATACTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2133)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAAGTATCCGACAGCGCCGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3706)	1	ATRIP_-_48488235.23-P1P2	ATRIP	P1P2	na	na	na	na	na	na	na	na	na
GTGCCGCCAGGCCGGAGAAC (SEQ ID NO: 561)	TATGAACCACTAAGGCGTccaccttgttgGTGCCGCCAGGCCGGAGAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2134)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTCTCCGGCCTGGCGGCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3707)	1	ATRIP_+_48488180.23-P1P2	ATRIP	P1P2	na	na	na	na	na	na	na	na	na

GGAGCCTCGGCCCAACAAAA (SEQ ID NO: 562)	TATGAACCACTAAGGCGTccaccttgttgGGAGCCTCGGCCCAACAAAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2135)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTTTGTTGGGCCGAGGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3708)	1		ATRX_+_77041684.23-P1P2	ATRX	P1P2	na	na	na	na	na	na	na	na	na
GGGCGCGGGGGAGCGCAGAT (SEQ ID NO: 563)	TATGAACCACTAAGGCGTccaccttgttgGGGCGCGGGGGAGCGCAGATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2136)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATCTGCGCTCCCCCGCGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3709)	1		ATRX_+_77041530.23-P1P2	ATRX	P1P2	na	na	na	na	na	na	na	na	na
GTGCCTGGGCCCAGACCGCT (SEQ ID NO: 564)	TATGAACCACTAAGGCGTccaccttgttgGTGCCTGGGCCCAGACCGCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2137)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGCGGTCTGGGCCCAGGCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3710)	1		ATRX_-_77041440.23-P1P2	ATRX	P1P2	na	na	na	na	na	na	na	na	na
GCGGCTTCCTAGTGAGTCGG (SEQ ID NO: 565)	TATGAACCACTAAGGCGTccaccttgttgGCGGCTTCCTAGTGAGTCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2138)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGACTCACTAGGAAGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3711)	1		BABAM1_-_17378297.23-P1P2	BABAM1	P1P2	na	na	na	na	na	na	na	na	na
GCTTCCGGCTGTAAAGATGG (SEQ ID NO: 566)	TATGAACCACTAAGGCGTccaccttgttgGCTTCCGGCTGTAAAGATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2139)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATCTTTACAGCCGGAAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3712)	1		BABAM1_-_17378278.23-P1P2	BABAM1	P1P2	na	na	na	na	na	na	na	na	na
GGCTACGGTGAGCCGAAGGT (SEQ ID NO: 567)	TATGAACCACTAAGGCGTccaccttgttgGGCTACGGTGAGCCGAAGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2140)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCTTCGGCTCACCGTAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3713)	1		BABAM1_-_17378340.23-P1P2	BABAM1	P1P2	na	na	na	na	na	na	na	na	na
GACAGGTCCAGTGAAGGCGC (SEQ ID NO: 568)	TATGAACCACTAAGGCGTccaccttgttgGACAGGTCCAGTGAAggcgcgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2141)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGCCTTCACTGGACCTGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3714)	1		BACH1_-_30671379.23-P1P2	BACH1	P1P2	na	na	na	na	na	na	na	na	na
GCGGCGGACAGGTCCAGTGA (SEQ ID NO: 569)	TATGAACCACTAAGGCGTccaccttgttgGCGGCGGACAGGTCCAGTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2142)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCACTGGACCTGTCCGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3715)	1		BACH1_-_30671373.23-P1P2	BACH1	P1P2	na	na	na	na	na	na	na	na	na
GGCGCCTCAGCTCTGGTGAG (SEQ ID NO: 570)	TATGAACCACTAAGGCGTccaccttgttgGGCGCCTCAGCTCTGGTGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2143)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCACCAGAGCTGAGGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3716)	1		BACH1_-_30671281.23-P1P2	BACH1	P1P2	na	na	na	na	na	na	na	na	na
GCGGCCCGCCGTCCCAGACG (SEQ ID NO: 571)	TATGAACCACTAAGGCGTccaccttgttgGCGGCCCGCCGTCCCAGACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2144)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTCTGGGACGGCGGGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3717)	1		BARD1_+_215674399.23-P1P2	BARD1	P1P2	na	na	na	na	na	na	na	na	na
GGGAAGCCTCGGGAAACCAC (SEQ ID NO: 572)	TATGAACCACTAAGGCGTccaccttgttgGGGAAGCCTCGGGAAACCACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2145)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTGGTTTCCCGAGGCTTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3718)	1		BARD1_-_215674358.23-P1P2	BARD1	P1P2	na	na	na	na	na	na	na	na	na
GTAATCGGCAGCCGAGGAAC (SEQ ID NO: 573)	TATGAACCACTAAGGCGTccaccttgttgGTAATCGGCAGCCGAGGAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2146)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTCCTCGGCTGCCGATTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3719)	1		BARD1_+_215674263.23-P1P2	BARD1	P1P2	na	na	na	na	na	na	na	na	na
GACTTCCCAGGAGCTCCGAG (SEQ ID NO: 574)	TATGAACCACTAAGGCGTccaccttgttgGACTTCCCAGGAGCTCCGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2147)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCGGAGCTCCTGGGAAGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3720)	1		BAZ1A_+_35344519.23-P1P2	BAZ1A	P1P2	na	na	na	na	na	na	na	na	na
GCCAGGAGCTCCGAGAGGCG (SEQ ID NO: 575)	TATGAACCACTAAGGCGTccaccttgttgGCCAGGAGCTCCGAGAGGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2148)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCTCTCGGAGCTCCTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3721)	1		BAZ1A_+_35344514.23-P1P2	BAZ1A	P1P2	na	na	na	na	na	na	na	na	na
GCCGAGAGGCGGGGAAGTGG (SEQ ID NO: 576)	TATGAACCACTAAGGCGTccaccttgttgGCCGAGAGGCGGGGAAGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2149)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACTTCCCCGCCTCTCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3722)	1		BAZ1A_+_35344505.23-P1P2	BAZ1A	P1P2	na	na	na	na	na	na	na	na	na
GCGGCCGCCGGGACCCGAAG (SEQ ID NO: 577)	TATGAACCACTAAGGCGTccaccttgttgGCGGCCGCCGGGACCCGAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2150)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTCGGGTCCCGGCGGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3723)	1		BAZ1B_+_72936507.23-P1P2	BAZ1B	P1P2	na	na	na	na	na	na	na	na	na
GGTCATGGAGCGGAACGCCA (SEQ ID NO: 578)	TATGAACCACTAAGGCGTccaccttgttgGGTCATGGAGCGGAACGCCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2151)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGCGTTCCGCTCCATGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3724)	1		BAZ1B_-_72936454.23-P1P2	BAZ1B	P1P2	na	na	na	na	na	na	na	na	na

GTCCTCCTGAGCAGCAC CGG (SEQ ID NO: 579)	TATGAAGCACTAAGGCGTcaacttgtggTCCTCCTCAGCAG TGCTGAGGAGCAACCAACAAGCTGGTGGTCATA (SEQ ID NO: 2152)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCCGGTGC 1	GAZTB_-_72935961.2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCAGGGAAGATGAAT TGTT (SEQ ID NO: 580)	TATGAACCACTAAGGCGTcaacttgtgGCGCAGGGAAGAAT CGTTCGCTGGTGGTCAACAAGGTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 2153)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCAATTCA 1	BGL11A_-_60780860.2 3-P1P2	BGL11A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTTGCGCGCAGACAT GGT (SEQ ID NO: 581)	TATGAACCACTAAGGCGTcaacttgtgGCTTCCGCGCGAG ACAATGGTgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2154)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCACCATGTC 1	BGL11A_-_60780410.2 3-P1P2	BGL11A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGCGTCGCAAGTTCAA GTG (SEQ ID NO: 582)	TATGAAGCACTAAGGCGTcaacttgtgGCTGAGCTCGCAAGTT CAAGTGgttaagagctaagtgcGCTAAAGTCTGCTGGTTGA (SEQ ID NO: 2155)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCACCTGAA 1	BGL11A_-_59785738.2 3-P1P2	BGL11A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAAGAAACCACTAAGGCGTcaacttgtgGAAAGAAACCACG GCATCGGgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2156)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCGGAGCTG GCTCGTTTCTTTCCACAAGGTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 2729)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCACCGGG 1	BGL11B_+_59797973.23-P1P2	BGL11B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTCGTCGTCTCGCCCG GGT (SEQ ID NO: 584)	TATGAACCACTAAGGCGTcaacttgtgGCTCTCGTCTCGTGG CCGCGGgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2157)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCACCGGG 1	BGL11B_-_59797895.2 3-P1P2	BGL11B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTTTCTTTCAGGGTCTG GT (SEQ ID NO: 585)	TATGAAGCACTAAGGCGTcaacttgtgGCTTTCTTTCAGGGG TCTGGTgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2158)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCACCAGAC 1	BGL11B_-_59798038.2 3-P1P2	BGL11B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGTTGCGCGGAGTGAGT GAG (SEQ ID NO: 586)	TATGAACCACTAAGGCGTcaacttgtgGAGTGCGAGTGA GTTGAGgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2159)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCTCACTCA 1	BGL11A_-_12245987.2 23-P1P2	BGL11A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAGCGCGGAGCAGAGC CGG (SEQ ID NO: 587)	TATGAACCACTAAGGCGTcaacttgtgGAAGCGCGCAGCAG GACCCGGgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2160)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCCCGAGTC 1	BGL12A_-_12245966.1 23-P1P2	BGL12A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCCCGGCCGCAACCAT GTC (SEQ ID NO: 588)	TATGAAGCACTAAGGCGTcaacttgtgGTCCCGGCCGGCAAA CCCATGTgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2161)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACGCAATCGG 1	BGL12A_-_12246004 23-P1P2	BGL12A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACCGAGCCGGCCGGACA TGG (SEQ ID NO: 589)	TATGAACCACTAAGGCGTcaacttgtgGACCGAGCCGGCCGC GACATgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2162)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCATGTC 1	BGL12B_-_72971912.2 3-P1P2	BGL12B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACTGCTGCTCCAAGAGA CCG (SEQ ID NO: 590)	TATGAACCACTAAGGCGTcaacttgtgGCACTGCTGCTCCAAG ACAACCGgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2163)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCGGTGTC 1	BGL12B_-_72971958.2 3-P1P2	BGL12B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGGACTGTGCCGGTCA GGT (SEQ ID NO: 591)	TATGAAGCACTAAGGCGTcaacttgtgGTGGGACTGTGCCGG GTCAAGGTgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2164)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACACCTGACC 1	BGL12B_+_72971738.23-P1P2	BGL12B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGTGCGGTGCCAAGT CCA (SEQ ID NO: 592)	TATGAACCACTAAGGCGTcaacttgtgGGAGGTGCGGTGCCA AACTCCAgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2165)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACGTGAGGTT 1	BGL12C_+_3695372.23-P1P2	BGL12C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCCAGGGAGCGGGACA CAG (SEQ ID NO: 593)	TATGAAGCCACTAAGGCGTcaacttgtgGGGCCAGGGAGCGG GACACAGgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2166)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACGTTGTGTCC 1	BGL12C_+_3695135.23-P1P2	BGL12C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCGCGGTGCCAAGTCCA GGG (SEQ ID NO: 594)	TATGAAGCACTAAGGCGTcaacttgtgGCGCGCGTGCCAAGT TCCAAGGGgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2167)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACCTCGTCGA 1	BGL12C_+_3695369.23-P1P2	BGL12C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGAGCGTCGCTAGCGCGT CGA (SEQ ID NO: 595)	TATGAACCACTAAGGCGTcaacttgtgGAGAGCGTCGCTAGC CCGTCGAgttaagagctaagtgcGCTAAAGTCTGCTGGTGA (SEQ ID NO: 2168)	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACTCGAGCC 1	BGLAF1_+_1389103946.23-P1P2	BGLAF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCCGGTCGAACGTAGCTCTA (SEQ ID NO: 596)	TATGAACCACTAAGGCGTcaacctgtgGCGGCGTCGAAGGTCCTTCGACCGCGCCAACAAGGTGGAGCGCCTTAGTCGTTCATA (SEQ ID NO: 2169)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACTAGAGCTA	35810395 23-P1P2	BCLAF1_+_135810395 23-P1P2	BCLAF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGGGAAATGTGCCGCT GTG (SEQ ID NO: 597)	TATGAACCACTAAGGCGTcaacctgtgGTCGGAAATGTGCCGCTCTTTCGCCGACCAAGAGGTCGAGCGCTTAGTCGTTCATA (SEQ ID NO: 2170)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACTAGAGCTA	35810894 23-P1P2	BCLAF1_+_135810894 23-P1P2	BCLAF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACCGCAGACGTCGTAGC TGGA (SEQ ID NO: 598)	TATGAACCACTAAGGCGTcaacctgtgGCACCGAAGTGGAGCGCTTAGTCGTTCATA (SEQ ID NO: 2171)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACTCCGGCTA	6945 23-P1P2	BLM_+_91286945 23-P1P2	BLM	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGCAGCGCAGACTCC TAG (SEQ ID NO: 599)	TATGAACCACTAAGGCGTcaacctgtgCCTCCGCAGCGCAGAGA CTCCTATGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2172)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACCTAGGAG TCTGCGTGCGAGCCAAGGTGGAGCGCTTAGTCGTTCATA	6949 23-P1P2	BLM_+_91286949 23-P1P2	BLM	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGCGGGAAGTTTCGGA TCC (SEQ ID NO: 600)	TATGAACCACTAAGGCGTcaacctgtgGGCCGCGGGAAGTTTCGGACTCCATATGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2173)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACGGACTCCA AACTTCCGCGCCGGCCAACAAGGTGGAGCGCTTAGTCGTTCATA	9_1260842 23-P1P2	BLM_+_91269842 23-P1P2	BLM	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGATCGGCGCGAGACA ATG (SEQ ID NO: 601)	TATGAACCACTAAGGCGTcaacctgtgGAGATCGGCGCGAGACA GACAATTgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2174)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACCATTGTCT CGCGCCGCGATCTCCAACAAGGTGGAGCGCTTAGTCGTTCATA	22610195 23-P1P2	BM11_-_22610195 23-P1P2	BM11	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACAGTGGAAACTGA CAC (SEQ ID NO: 602)	TATGAACCACTAAGGCGTcaacctgtgCGCAGAGTGGAAA CTGACACTGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2175)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACGTGTCAGT TTCGCACTGTCGCCAACAAGGTGGAGCGCTTAGTCGTTCATA	10019 23-P1P2	BM11_+_22610019 23-P1P2	BM11	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGAGACAATGGGGCA TGT (SEQ ID NO: 603)	TATGAACCACTAAGGCGTcaacctgtgGGGCGAGACAATG GCGGATGTgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2176)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACCATCCCCC ATTGTCTCGCCCCAAGCAAGGTGGAGCGCTTAGTCGTTCATA	22611202 23-P1P2	BM11_-_22611202 23-P1P2	BM11	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACAGCCTGTCCCCCCG TCC (SEQ ID NO: 604)	TATGAACCACTAAGGCGTcaacctgtgGCACAGCCCTGTCTCC CCCGTCCgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2177)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACGGGAGGGG TGGACAGGCTGTGGCCAACAAGGTGGAGCGCTTAGTCGTTCATA	41277370 23-P1P2	BRCA1_-_41277370 23-P1P2	BRCA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAAGGCCTTCCTGAGCG CAG (SEQ ID NO: 605)	TATGAAGCACTAAGGCGTcaacctgtgGGAAAGGCCTTCCTGAGCG AGGCGCAGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2178)	TCAGCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACCTGCGCT CAGGGAGGCCTTCCCCAACAAGGTGGAGCGCTTAGTCGTTCATA	41277228 23-P1P2	BRCA1_-_41277228 23-P1P2	BRCA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGCTGAGACTTCCTCTG GAC (SEQ ID NO: 606)	TATGAACCACTAAGGCGTcaacctgtgGTCGCTGAGACTTCCTCTGCTTCCTGGTCAGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2179)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAAGGTCCGAGG AAGTCTCAGCGGACCAACAAGGTGGAGCGCTTAGTCGTTCATA	12773817 23-P1P2	BRCA1_+_412773817 23-P1P2	BRCA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACACCCGAGGCGCGAG CAG (SEQ ID NO: 607)	TATGAACCACTAAGGCGTcaacctgtgGGACACCCGAGGCGCGAG GCACGCAGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2180)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACCTGCTGC GCCTCCGGGTGTCCCAACAAGGTGGAGCGCTTAGTCGTTCATA	28898673 23-P1P2	BRCA2_+_328898673 23-P1P2	BRCA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGGGGACAGATTTGT GAC (SEQ ID NO: 608)	TATGAAGCACTAAGGCGTcaacctgtgCGACAGGGGACAGACAT TTGTCGACgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2181)	TCAGCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACGTCACACAA TCTGTCCCCTCCCCAACAAGGTGGACGCCTTAGTCGTTCATA	32869758 23-P1P2	BRCA2_-_328869758 23-P1P2	BRCA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGTCAGGTTCTCTGAA TGAAAC (gttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 609)	TATGAACCACTAAGGCGTcaacctgtgGGCGTCAGGTTCTCTGAA AACTCGGGCGCCGAACAAGGTGCGACGCCTTAGTCGTTCATA (SEQ ID NO: 2182)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACGATTTCAG AAGCTCCGGGCCCGAACAAGGTGCGACGCCTTAGTCGTTCATA	32869838 23-P1P2	BRCA2_-_328869838 23-P1P2	BRCA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGAGGTCGGACGTGA GAA (SEQ ID NO: 610)	TATGAACCACTAAGGCGTcaacctgtgGCCCGAGGTCGGACGTGA GTGAGAAgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2183)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACTTCTCACG TCCGCCTCCGGGCCAACAAGGTGGAGCGCTTAGTCGTTCATA	15428878 23-P1P2	BRCC3_-_154288781 23-P1P2	BRCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGAAGCGGTCCGGGCCA AGA (SEQ ID NO: 611)	TATGAAGCACTAAGGCGTcaacctgtgCGAGAAGTGGTCCGGG CCAGCGAAGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2184)	TCAGCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACTCTTGGTC CGAGCCGCTTCTCCCAACAAGGTGGAGCGCTTAGTCGTTCATA	15429366 23-P1P2	BRCC3_-_154293666 23-P1P2	BRCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAACCAAGTGTGGCCCG AGG (SEQ ID NO: 612)	TATGAACCACTAAGGCGTcaacctgtgGGAACCAAGTGTGGCCCG AGGCCACGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2185)	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACGCCTGGTG CGCCACTTGGTGCCCAACAAGGTGGAGCGCTTAGTCGTTCATA	15429379 23-P1P2	BRCC3_-_154293779 23-P1P2	BRCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCCCCTCCTCACCCTCGTAG (SEQ ID NO: 613)	TATGAACCACTAAGGCGTccaccttgttgGccccTCCTCACCCTCGTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2186)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTACGAGGGTGAGGAGGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3759)	1		BRD7_-_50402646.23-P1P2	BRD7	P1P2	na	na	na	na	na	na	na	na	na
GGATGCCCCTCTCGAGAAGA (SEQ ID NO: 614)	TATGAACCACTAAGGCGTccaccttgttgGgATGCCCCTCTCGAGAAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2187)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTTCTCGAGAGGGGCATCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3760)	1		BRD7_-_50402654.23-P1P2	BRD7	P1P2	na	na	na	na	na	na	na	na	na
GGGGCCCGCCGCCCGCACCC (SEQ ID NO: 615)	TATGAACCACTAAGGCGTccaccttgttgGgggcccgccgcccgcacccgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2188)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGTGCGGGCGGCGGGCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3761)	1		BRD7_-_50402623.23-P1P2	BRD7	P1P2	na	na	na	na	na	na	na	na	na
GCCAAGATGGCGACGGGAAC (SEQ ID NO: 616)	TATGAACCACTAAGGCGTccaccttgttgGCCAAGATGGCGACGGGAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2189)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTCCCGTCGCCATCTTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3762)	1		BRD8_+_137514286.23-P1P2	BRD8	P1P2	na	na	na	na	na	na	na	na	na
GCTCAGGGTTGGAGACTTCG (SEQ ID NO: 617)	TATGAACCACTAAGGCGTccaccttgttgGCTCAGGGTTGGAGACTTCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2190)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGAAGTCTCCAACCCTGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3763)	1		BRD8_+_137514308.23-P1P2	BRD8	P1P2	na	na	na	na	na	na	na	na	na
GTCGGGGCCAAGATGGCGAC (SEQ ID NO: 618)	TATGAACCACTAAGGCGTccaccttgttgGTCGGGGCCAAGATGGCGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2191)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCGCCATCTTGGCCCCGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3764)	1		BRD8_+_137514292.23-P1P2	BRD8	P1P2	na	na	na	na	na	na	na	na	na
GGCCGCCGCCTCACCCTCGT (SEQ ID NO: 619)	TATGAACCACTAAGGCGTccaccttgttgGGCCGCCGCCTCACCCTCGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2192)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGAGGGTGAGGCGGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3765)	1		BRD9_-_892728.23-P1P2	BRD9	P1P2	TRIP13_-_892728.23-P1P2	TRIP13	P1P2	na	na	na	na	na	na
GTCGTCCTACGAGGGTGAGG (SEQ ID NO: 620)	TATGAACCACTAAGGCGTccaccttgttgGTCGTCCTACGAGGGTGAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2193)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTCACCCTCGTAGGACGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3766)	1		BRD9_+_892711.23-P1P2	BRD9	P1P2	TRIP13_+_892711.23-P1P2	TRIP13	P1P2	na	na	na	na	na	na
GCCAGCGAGCTCGGCAACCT (SEQ ID NO: 621)	TATGAACCACTAAGGCGTccaccttgttgGCCAGCGAGCTCGGCAACCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2194)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGGTTGCCGAGCTCGCTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3767)	1		BRD9_+_892880.23-P1P2	BRD9	P1P2	na	na	na	na	na	na	na	na	na
GACTTAGTCCGATCGGTAAC (SEQ ID NO: 622)	TATGAACCACTAAGGCGTccaccttgttgGACTTAGTCCGATCGGTAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2195)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTACCGATCGGACTAAGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3768)	1		BRE_+_28113682.23-P1P2	BRE	P1P2	na	na	na	na	na	na	na	na	na
GAGGTCGGTGCGTCTGTCGG (SEQ ID NO: 623)	TATGAACCACTAAGGCGTccaccttgttgGAGGTCGGTGCGTCTGTCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2196)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGACAGACGCACCGACCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3769)	1		BRE_-_28113650.23-P1P2	BRE	P1P2	na	na	na	na	na	na	na	na	na
GGACTCGCCTCCTCTTCCTA (SEQ ID NO: 624)	TATGAACCACTAAGGCGTccaccttgttgGGACTCGCCTCCTCTTCCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2197)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGGAAGAGGAGGCGAGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3770)	1		BRE_-_28113815.23-P1P2	BRE	P1P2	na	na	na	na	na	na	na	na	na
GCGGAGGGAAGTCCACGGAA (SEQ ID NO: 625)	TATGAACCACTAAGGCGTccaccttgttgGCGGAGGGAAGTCCACGGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2198)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCCGTGGACTTCCCTCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3771)	1		BRIP1_+_59940616.23-P1P2	BRIP1	P1P2	na	na	na	na	na	na	na	na	na
GGGTCTGGTGGGTCGAGGAA (SEQ ID NO: 626)	TATGAACCACTAAGGCGTccaccttgttgGGGTCTGGTGGGTCGAGGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2199)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCCTCGACCCACCAGACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3772)	1		BRIP1_+_59940833.23-P1P2	BRIP1	P1P2	na	na	na	na	na	na	na	na	na
GTGAGGGGTCTGGTGGGTCG (SEQ ID NO: 627)	TATGAACCACTAAGGCGTccaccttgttgGTGAGGGGTCTGGTGGGTCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2200)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGACCCACCAGACCCCTCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3773)	1		BRIP1_+_59940838.23-P1P2	BRIP1	P1P2	na	na	na	na	na	na	na	na	na
GCTGCGCGGGGTATTCGAAT (SEQ ID NO: 628)	TATGAACCACTAAGGCGTccaccttgttgGCTGCGCGGGGTATTCGAATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2201)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATTCGAATACCCCGCGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3774)	1		BUB1_+_111435629.23-P1P2	BUB1	P1P2	na	na	na	na	na	na	na	na	na
GGAATGTGTCCCACCCAGCA (SEQ ID NO: 629)	TATGAACCACTAAGGCGTccaccttgttgGGAATGTGTCCCACCCAGCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2202)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGCTGGGTGGGACACATTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3775)	1		BUB1_+_111435409.23-P1P2	BUB1	P1P2	na	na	na	na	na	na	na	na	na

GGGTTGCCCAACCCAACC AGA (SEQ ID NO: 630)	TATGAAGCACTAAGGCGTcaacctgtgGGCTTGCCCAACCC CAACCACAgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2203)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAAACTCTGGTG G9GTGGGGAAGCCCAACAAGGTCGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3776)	BUB1_111436537 23-P1P2	BUB1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGTATTCGAGCCTTG GAC (SEQ ID NO: 631)	TATGAACCACTAAGGCGTcaacctgtgGCTGTATTCGAGCC GTTGGAAgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2204)	T:CAACGAGCAGACTTAGCGTCTTAGCAGCTTAAACTCTGG GGCCGAATACAGCCAACAAGGTGGACGGCTTAGTGGTTCGATA (SEQ ID NO: 3777)	C19orf40_33463237.2 3-P1P2	C19orf40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTTGCTGCCGGCTGT ATT (SEQ ID NO: 632)	TATGAACCACTAAGGCGTcaacctgtgGGTTGCCTGCCGGG GTGTATTgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2205)	T:CAACGAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACAATACAGGC G9GCAGTGAACCCCAACAAGGTGGACGGCTTAGTGGTTCGATA (SEQ ID NO: 3778)	C19orf40_33463225.2 3-P1P2	C19orf40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCTCTACACTGGGAACG GTG (SEQ ID NO: 633)	TATGAACCACTAAGGCGTcaacctgtgGTCTCTACACTGGG GACGGTTGgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2206)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAAACTCCACGGTG CCAGTCTAGAGACGCAAGGTGGACGGCTTAGTGGTTCGATA (SEQ ID NO: 3779)	C19orf40_33463205.2 3-P1P2	C19orf40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCATTTAGCCACATGGG AA (SEQ ID NO: 634)	TATGAACCACTAAGGCGTcaacctgtgGCCATTTAGCCACA TGGGAAgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2207)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAAACTTCCCATG T9GGTAAATGGGCAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3780)	C1orf86 + 213013423-P1P2	C1orf86	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTTGACGCTCACTTCC GCC (SEQ ID NO: 635)	TATGAACCACTAAGGCGTcaacctgtgGgCTTGCGCTCACT TCCGCGgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2208)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACGGGGGAA GTGAGCGCAAGCCCAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3781)	C1orf86 + 213019923-P1P2	C1orf86	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGCTCAACCCACAG CCG (SEQ ID NO: 636)	TATGAACCACTAAGGCGTcaacctgtgCgggcgctCAACCCC CAGCCGgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2209)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAAACTTCCTCGCTCGC GGTTGAGCCGCGCCCAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3782)	C1orf86_-21295169.23-P1P2	C1orf86	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGCGAGGCTGGGTTAC GTG (SEQ ID NO: 637)	TATGAACCACTAAGGCGTcaacctgtgGCGGCGAGGCTGGG TTACGTGgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2210)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACCACGTAAC CCAGCCTCGGCGCCAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3783)	CABIN1_24407927.23-P1P2	CABIN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAATGGGTTCCGGGCCT CCG (SEQ ID NO: 638)	TATGAACCACTAAGGCGTcaacctgtgGGAATGGGTTCCGG GCCTCCGgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2211)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACCGGGAGGC CGGGAACCCCATTCCCAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3784)	ICABIN1_24409662.23-P1P2	CABIN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACCCTCAGGGTGCCTAG AAT (SEQ ID NO: 639)	TATGAAGCACTAAGGCGTcaacctgtgGGCCCTCAGGGTG CCTAGAATgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2212)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACATTCGTAGGC CACCTCAGGGCCCAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3785)	CABIN1_24408047.23-P1P2	CABIN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGGGGCGTGTGTAAAC AAA (SEQ ID NO: 640)	TATGAACCACTAAGGCGTcaacctgtgGCCGGCGGGGCGTGT TAAACAAAgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2213)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACGTTTGTTTA CACCGCCGCGGGCGACACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3786)	CBX2_+_77752406.23-P1P2	CBX2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGGGTCCGGGGCGGG TGAC (SEQ ID NO: 641)	TATGAACCACTAAGGCGTcaacctgtgGgggggtccgggggcgggg acgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2214)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACGGTCACGCC GCCGGGAACGCGCCCCAAACAAGGTGGACGGCTTAGTGGTTCAT A (SEQ ID NO: 3787)	CBX2_-77761997.23-P1P2	CBX2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTCCGCGGGGGTGAC TGG (SEQ ID NO: 642)	TATGAAGCACTAAGGCGTcaacctgtgCgggtccgggggcgggtgaag gattaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2215)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACGCCAGTCAC CCCGCCCGGAGCCCCAAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3788)	CBX2_-77752003.23-P1P2	CBX2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCATGGAGCTGCCCAGC TGT (SEQ ID NO: 643)	TATGAACCACTAAGGCGTcaacctgtgGCCATGGAGCTGC CCAGCTGTgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2216)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACGACAGGTG GCAGCGTCCATGGCGCAAGAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3789)	CBX4_+_77813030.23-P1P2	CBX4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAACAACTGAGCCCGGCC CCG (SEQ ID NO: 644)	TATGAACCACTAAGGCGTcaacctgtgGGAACAACTGAGCC CCGGCGCGgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2217)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACCGGGCCCG9 GGCTCAGTTGTCGCCAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3790)	CBX4_-77813253.23-P1P2	CBX4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGTCGCCCAAGTTAAG TCG (SEQ ID NO: 645)	TATGAACCACTAAGGCGTcaacctgtgGTGGTCGCCCAAG TTAAGTCGgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2218)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACCCACTTAC TTGGGCGGACCCACAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3791)	CBX4_+_77812728.23-P1P2	CBX4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CAGGGAGTATTATGGGC TGT (SEQ ID NO: 646)	TATGAACCACTAAGGCGTcaacctgtgCAGGGAGTATTATG GCTGTgttaagagttaagtgtGGTAAAGTTCTGCTGGTGA (SEQ ID NO: 2219)	T:CAGCAGCAGACTTAGCGTCTTAGCAGCTTAGCTCTTAAACGACAGGCC ATAATACTCCCTCGCAACAAGGTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 3792)	CBX6_+_38268059.23-P1P2	CBX6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCGTGCGCCGCGGCCG TG (SEQ ID NO: 647)	TATGAACCACTAAGGCGTcaacctgtgGCGTaagcaggcgg cgtgattaagcgtcaagcgtCGTAAATGTCGTCGTGTGA (SEQ ID NO: 2220)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCACCGGC 1 CGCGGGCGCACGG (CAACAAGGTGGAGCGCTTAGTGGTTCAT A (SEQ ID NO: 3793)	CBX6_-_39268249.2 3-P1P2	CBX6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGCGCAGGGAGTAT TAT (SEQ ID NO: 648)	TATGAACCACTAAGGCGTcaacctgtgGaggtgcgcAAGGAGT ATTATgttaagcgtaagcgtGCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2221)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCATAATACT 1 CCCTGCGCCCGCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3794)	CBX6_+_382882216.23-P1P2	CBX6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGTGCGCCGCCGA CGA (SEQ ID NO: 649)	TATGAACCACTAAGGCGTcaacctgtgGcgtgtgtgtgcgaGA CGAgttaagcgtaagcgtGCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2222)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCCGTCCG 1 GGCGCGCAGCGCGCAACAAGGTGGAGCGCTTAGTGGTTCAT A (SEQ ID NO: 3795)	CBX7_+_39548623-P1P2	CBX7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGTCAGCCATCGGCG AGC (SEQ ID NO: 650)	TATGAACCACTAAGGCGTcaacctgtgCGTGTCAGCCATG 1 GGTCGAGCgttaagcgtaagcgtgCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2223)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGCTCCGCC 1 GATGGCTGACAGCCAAGAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3796)	CBX7_+_39548623-P1P2	CBX7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACGGGAAGGAGACCCA CTG (SEQ ID NO: 651)	TATGAACCACTAAGGCGTcaacctgtgGACGGAGAGGAC CCCAACTGgttaagcgtaagcgtgCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2224)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCAGTGGG 1 GTCCCTCCCGCGTCCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3797)	CBX7_-_39548350.2 3-P1P2	CBX7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTCGCCGCGTTCCCC CCT (SEQ ID NO: 652)	TATGAACCACTAAGGCGTcaacctgtgGGACTTCCGCCGCTT CCCCCCTgttaagcgtaagcgtgCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2225)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACAGGGGGG 1 AACGCGCGGAGTCCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3798)	ICBX8_-_77779921.23-P1P2	CBX8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGAAGCGGCGACTCA ACA (SEQ ID NO: 653)	TATGAACCACTAAGGCGTcaacctgtgCGGCGAAGCGGCGA 1 CGCCGCTTCCCCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3690)	CBX8_+_7777079323.23-P1P2	CBX8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCTCCTGGAAGCGCGC CAA (SEQ ID NO: 654)	TATGAACCACTAAGGCGTcaacctgtgGTGCTCCTGGAAG 1 CTTCCAGGAGCAGCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3801)	CBX8_+_7777079315.23-P1P2	CBX8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGCTCTGTAGGAGGA GAG (SEQ ID NO: 655)	TATGAACCACTAAGGCGTcaacctgtgGAGCTGTCTGTAGG 1 CTACAGAGCAGCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 2228)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCCTCTCTC 1 CCCCTCCCGCGTCCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3802)	CCDC58_+_122102127.23-P1P2	CCDC58	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGCAATCAGCGTCGGAG GAA (SEQ ID NO: 656)	TATGAAGCACTAAGGCGTcaacctgtgCGCGCAATCAGCGC 1 GGAGCAAgttaagcgtaagcgtgCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2229)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTTGCTCCG 1 CGAGTGATTGGCCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3692)	CCDC58_-_122102121.23-P1P2	CCDC58	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGCGCGAGTTCCACGGT GAT (SEQ ID NO: 657)	TATGAACCACTAAGGCGTcaacctgtgGTCGCCGAGTTCC 1 AGGTCGATgttaagcgtaagcgtgCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2230)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACATCACCTG 1 GAAGTCCGCGCAACCAACAAGGTGGACGCTTAGTGGTTCATA (SEQ ID NO: 3803)	CCDC58_+_122102012.23-P1P2	CCDC58	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGATAACGACGGGGAA GAG (SEQ ID NO: 658)	TATGAACCACTAAGGCGTcaacctgtgGAAGATAACGACG 1 GGAAGAGgttaagcgtaagcgtgCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2231)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCTCTTCCC 1 GTCGTTATCTTCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3804)	CCNA1_-_37036972.23-P1P2	CCNA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGATAAGAAGCATAACG AC (SEQ ID NO: 659)	TATGAACCACTAAGGCGTcaacctgtgCGACATAGAAGAG 1 AACGAGgttaagcgtaagcgtgCGTAAAGTGTCGTCGTGTGA (SEQ ID NO: 2232)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGTCGTTAT 1 CTTTCTATCTTGCCAACAAGGTGGACGCTTAGTGGTTCATA (SEQ ID NO: 3805)	CCNA1_-_37036954.23-P1P2	CCNA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCTCAGCGTGGGAACGA AGG (SEQ ID NO: 660)	TATGAACCACTAAGGCGTcaacctgtgGTCTCAGCGTGGGA 1 ACGAACGgttaagcgtaagcgtgCGTAAAGTCTCGTCGTGTGA (SEQ ID NO: 2233)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGCGTCCTT 1 CGCGAGCTGAGACGAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3806)	CCNA1_-_37036944.23-P1P2	CCNA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGCAGCAGCGGGTGT AGG (SEQ ID NO: 661)	TATGAACCACTAAGGCGTcaacctgtgGGCGCAGCAGCGG 1 GTGTACGGgttaagcgtaagcgtgCGTAAAGTCTCGTCGTGTGA (SEQ ID NO: 2234)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCTACACG 1 GCCGCTCGCGGCGCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3807)	CCNE1_-_30302934.23-P1P2	CCNE1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCGCGCGTCGAGCGCG GTGT (SEQ ID NO: 662)	TATGAACCACTAAGGCGTcaacctgtgCGCGCGTCGACGCG 1 GCGGGTGTgttaagcgtaagcgtgCGTAAAGTCTCGTCGTGTGA (SEQ ID NO: 2235)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACACACCGC 1 GCTGCGCGCGGCGCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3808)	CCNE1_-_30302393.23-P1P2	CCNE1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCGCGCGCGGACAAGAA CCC (SEQ ID NO: 663)	TATGAACCACTAAGGCGTcaacctgtgGTCCGCGCGGGAC 1 AAGAAGCCgttaagcgtaagcgtgCGTAAAGTCTCGTCGTGTGA (SEQ ID NO: 2236)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGGGGTGTT 1 GTCCGCGCGGCCGACCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 3809)	CCNE1_-_30303059.23-P1P2	CCNE1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GAAGATCTGGCGCGAAGC CTA (SEQ ID NO: 664)	TATGAAGCACTAAGGCGTcaacttgtgGAAGATCTCGCGG AAGCCTTAgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2237)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTTAGGCTTC CGCGAGATCTCTCCAAACGAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3810)	CCNH_-89708718.2 3-P1P2	CCNH	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTTCCGCGAGATCTTC TGA (SEQ ID NO: 665)	TATGAACCACTAAGGCGTcaacttgtgGGCTTCCGCGAGA TCTTCCGgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2238)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCGAAGA TCTTCGCCTCGAAGCGCCAACAAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3811)	CCNH_+_86708692.23-P1P2	CCNH	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCGCGAGATCTTCGTG TGG (SEQ ID NO: 666)	TATGAACCACTAAGGCGTcaacttgtgGTCCGAGATCTTCG TGGttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2239)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCGACGGA AGATCTCGCGAGGAAGCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3812)	CCNH_+_86708689.23-P1P2	CCNH	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAACCGTGAACCCGAACC CAT (SEQ ID NO: 667)	TATGAACCACTAAGGCGTcaacttgtgCGAACCCGTGAACCC GAACCATgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2240)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCATGGTTCG GAACCATGGTCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3813)	CDC40_+_110650678.23-P1P2	CDC40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAACCGGTACCTAACGA AG (SEQ ID NO: 668)	TATGAACCACTAAGGCGTcaacttgtgGTAACCGGTACCT AACGAAGgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2241)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCTTGGTTAG GTAACCGGTCTTACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3814)	CDC40_+_110651657.23-P1P2	CDC40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTACCTTAACTGCCACC TC (SEQ ID NO: 669)	TATGAACCACTAAGGCGTcaacttgtgGTTACCTTAACTGCC CACCTCgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2242)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGAGGTGG CACCTCTTAAGGTTAACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3815)	CDC40_+_110650619.23-P1P2	CDC40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAATTGACGCGAGAGAGG AGG (SEQ ID NO: 670)	TATGAACCACTAAGGCGTcaacttgtgGAATTGACGCGAG AGGAGGgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2243)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGTCGCGCT CTCGCGTCAATTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3816)	CDK1_-_625638264.23-P1P2	CDK1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAACGGCCGAGGGCCTC GGA (SEQ ID NO: 671)	TATGAACCACTAAGGCGTcaacttgtgGGAACGGCCGAGGG CCCTCGGGAgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2244)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTCCGAGGG CCCTCGGGCGGTCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3817)	CDK1_-_625636382.23-P1P2	CDK1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTACCGGCGCTTATTATTC CG (SEQ ID NO: 672)	TATGAACCACTAAGGCGTcaacttgtgGTACCGGCGCTTATT ATTCCGgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2245)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGAATAA TAAGGCGGGTACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3818)	CDK1_+_625638369.23-P1P2	CDK1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCTCTTCGAACGACGG CGA (SEQ ID NO: 673)	TATGAAGCACTAAGGCGTcaacttgtgGCTCTCTTCGAACG ACGGGGAgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2246)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCCGTTC CTTCCATGAGAGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3819)	CDK7_-_58503885.23-P1P2	CDK7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCCGAGTCCACGCGA AAA (SEQ ID NO: 674)	TATGAACCACTAAGGCGTcaacttgtgGGCCCGAGTCCACG CGAAAAgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2247)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGTTTTTCGGGC TCGGGCTCGGGCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3820)	CDK7_+_36530767.23-P1P2	CDK7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGGCGATACTAAAG CGA (SEQ ID NO: 675)	TATGAACCACTAAGGCGTcaacttgtgGGGCGGGCGATACT AAAGCGAgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2248)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCTTTA GTATCGCCCGCCCGAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3821)	CDK7_-_58503885.23-P1P2	CDK7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTCCTCGGAACGCCAT GGT (SEQ ID NO: 676)	TATGAAGCACTAAGGCGTcaacttgtgGGCTCCTCGGAAG CCATGGTgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2249)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGACCATGG CTTCCGAGGAGGCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3822)	CDYL_+_47776111.23-P1	CDYL	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGCGGCGCGTCCGG ATCG (SEQ ID NO: 677)	TATGAACCACTAAGGCGTcaacttgtgGGGGGCGGCGCGTC CGGATCGgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2250)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGGGATCCG GGCGCGCCGCGCCCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ A (SEQ ID NO: 3823)	CDYL_+_47798680.23-P1	CDYL	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCTGCACTACACCGGGA GAG (SEQ ID NO: 678)	TATGAACCACTAAGGCGTcaacttgtgGTCTGCACTACACC GGGAGAGgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2251)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCTCTCCCG GTCTGCACTACACCGGGAGAGCGACGCCTTAGTGGTTCATA (SEQ A (SEQ ID NO: 3824)	CDYL_-_47778677.23-P1	CDYL	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACAACGAGCGGCGACGAG GTC (SEQ ID NO: 679)	TATGAAGCACTAAGGCGTcaacttgtgGACAACGAGCGGC GACGAGGTCgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2252)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGTCGCTGC TCCGGTTGTTGTCGCAACGAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3825)	CDYL_-_47734412.23-P2	CDYL	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCAAACAATTGATCATC GA (SEQ ID NO: 680)	TATGAACCACTAAGGCGTcaacttgtgGCCAAACAATTGAT CATCGAgttaagagctaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 2253)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGTCGATGAT CAATTGTTTGGCGCAACGAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3826)	CDYL_+_47733351.23-P2	CDYL	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGTTATCATTCGTGCGAG AT (SEQ ID NO: 681)	TATGAAGCACTAAGGCGTcaacttgtgGGTTATCATTCGTCTC GACAATGtatcaagccaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2254)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACATCTCGG3 4773285 23-P2	CDYL	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGTCGGTTGCCTAACA AC (SEQ ID NO: 682)	TATGAACCACTAAGGCGTcaacttgtgGACGTCGGTTGCC GCAACCGAGGTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2255)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACGTTGTTAG 1999272 23-P1	CETN2_+_15	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGAGTCAGTGTACA CGT (SEQ ID NO: 683)	TATGAACCACTAAGGCGTcaacttgtgGGCCGAGTCAGTGT TACAGGTgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2256)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACACGTGTAG 1999257 23-P1	CETN2_+_15	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGTCGCGCCCCAACGCC CAG (SEQ ID NO: 684)	TATGAACCACTAAGGCGTcaacttgtgGTCGTCGCGCCCCA TGGGGGCGCCAACAAGGTCGACGCCTTAGTGGTTCAT A (SEQ ID NO: 2257)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCTGGGGCC	CETN2_+_15	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGGGGTATAGAGAGGTG AATA (SEQ ID NO: 685)	TATGAACCACTAAGGCGTcaacttgtgGAAGGGGTATAGAG CGTGATAgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2258)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTATCACCT CTCTAGCCCTTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3831)	CETN2_-_151998683 23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGAAAAGGAAGGGGCA GTA (SEQ ID NO: 686)	TATGAACCACTAAGGCGTcaacttgtgGGAGGAAAAGGAAG GGCAAGTAgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2259)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTACTGCCC TTCCTTTCCTCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3832)	CETN2_-_151998724 23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTCAGTTCTTTAACATC TA (SEQ ID NO: 687)	TATGAACCACTAAGGCGTcaacttgtgGTTCAGTTCTTTAA CATCTAgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2260)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTAGATGTT AAAGAACTGAACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3833)	CETN2_-_151998167 23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGGAGAGGCGAAGA GCAG (SEQ ID NO: 688)	TATGAACCACTAAGGCGTcaacttgtgGGGCGGGAGAGGCG caagcttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2261)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTGCTCTT 4402698 23-P1P2	CHAF1A_-_	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCAGCGTCGACGTCCCT CTC (SEQ ID NO: 689)	TATGAACCACTAAGGCGTcaacttgtgGGGCAGCGTCGACG GGTCGAGCTGGCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2262)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACGAGAGGA 4402728 23-P1P2	CHAF1A_+_	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCGCGGGGCTCCCCTC AGG (SEQ ID NO: 690)	TATGAAGCACTAAGGCGTcaacttgtgCGGCGGGCTCCCC CCTCGAggttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2263)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCCTGAGGG 4402599 23-P1P2	CHAF1A_+_	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGAGGGAAGCCGTACCT GGA (SEQ ID NO: 691)	TATGAACCACTAAGGCGTcaacttgtgGCGAGGGAAGCCGT AGCCTGGAgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2264)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAAGTCCAGGTA 3775741 23-P1P2	CHAF1B_+_	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAGGGACCCGGACAGT CAG (SEQ ID NO: 692)	TATGAACCACTAAGGCGTcaacttgtgGGAGGGACCCGGGA CAGTCAGgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2265)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCTGACTGT 3775724 23-P1P2	CHAF1B_+_	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCCGGCGTCGCCTCCAG GTA (SEQ ID NO: 693)	TATGAAGCACTAAGGCGTcaacttgtgCGTCCGGCGTCGCCT CCAGGTAgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2266)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTACCTGGA 3775775 42 3-P1P2	CHAF1B_-_	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACTGAGGCGCCGAGCCG CCG (SEQ ID NO: 694)	TATGAACCACTAAGGCGTcaacttgtgGACTGAGGCGCCGA CCGCGACGgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2267)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACGCGTCGGC 1 CCGGCCCTCAGTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3840)	CHD1L_-_146714437 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGCGCCTGGCTTGTTACC TG (SEQ ID NO: 695)	TATGAACCACTAAGGCGTcaacttgtgGAGCGCCTGGCTT GTTACTGgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2268)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCAGTAAGA AGCCCAGCGGCGTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3841)	CHD1L_-_146714469 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAAGTCCGGAGTAAGAA GCC (SEQ ID NO: 696)	TATGAAGCACTAAGGCGTcaacttgtgGTGAAGTCCGGAGT AAGAAGCCgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2269)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACGTGGTTA 6714392 23-P1P2	CHD1L_+_14	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CAACGACGTTGGGCTCTCGA GTA (SEQ ID NO: 697)	TATGAACCACTAAGGCGTcaacttgtgCAACGACGTTGGGCT CTCAGTAgttaagcgtcaagcgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 2270)	TCAGCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTACTGTGAG 43780 23-P1	CHD2_+_934	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGCAGAGCGCGCTCTCTCTA (SEQ ID NO: 698)	TATGAACCACTAAGGCGTccaccttgttgGGCAGAGCGCGCTCTCTCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2271)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGAGAGAGCGCGCTCTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3844)	1	CHD2_+_93443587.23-P1	CHD2	P1	na	na	na	na	na	na	na	na	na
GGTTGCGTAAAGTGAGGCAA (SEQ ID NO: 699)	TATGAACCACTAAGGCGTccaccttgttgGGTTGCGTAAAGTGAGGCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2272)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGCCTCACTTTACGCAACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3845)	1	CHD2_+_93443615.23-P1	CHD2	P1	na	na	na	na	na	na	na	na	na
GCAACATGGCCTCTTCGCTG (SEQ ID NO: 700)	TATGAACCACTAAGGCGTccaccttgttgGCAACATGGCCTCTTCGCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2273)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGCGAAGAGGCCATGTTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3846)	1	CHD2_+_93425965.23-P2	CHD2	P2	na	na	na	na	na	na	na	na	na
GGAGGGAGTAAAGCCTAGAG (SEQ ID NO: 701)	TATGAACCACTAAGGCGTccaccttgttgGGAGGGAGTAAAGCCTAGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2274)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCTAGGCTTTACTCCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3847)	1	CHD2_-_93426037.23-P2	CHD2	P2	na	na	na	na	na	na	na	na	na
GTAAAGCCTAGAGAGGAGAC (SEQ ID NO: 702)	TATGAACCACTAAGGCGTccaccttgttgGTAAAGCCTAGAGAGGAGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2275)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCTCCTCTCTAGGCTTTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3848)	1	CHD2_-_93426044.23-P2	CHD2	P2	na	na	na	na	na	na	na	na	na
GAGGATGGCTTCCCCTCTGA (SEQ ID NO: 703)	TATGAACCACTAAGGCGTccaccttgttgGAGGATGGCTTCCCCTCTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2276)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCAGAGGGGAAGCCATCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3849)	1	CHD3_-_7788142.23-P1P2	CHD3	P1P2	na	na	na	na	na	na	na	na	na
GGCTTCCCCTCTGAGGGACG (SEQ ID NO: 704)	TATGAACCACTAAGGCGTccaccttgttgGGCTTCCCCTCTGAGggacggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2277)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTCCCTCAGAGGGGAAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3850)	1	CHD3_-_7788148.23-P1P2	CHD3	P1P2	na	na	na	na	na	na	na	na	na
GTCCTCCTCGTCCCTCAGAG (SEQ ID NO: 705)	TATGAACCACTAAGGCGTccaccttgttgGtcctcctcgtccCTCAGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2278)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCTGAGGGACGAGGAGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3851)	1	CHD3_+_7788131.23-P1P2	CHD3	P1P2	na	na	na	na	na	na	na	na	na
GCCGGACCCAGCGCGACCTG (SEQ ID NO: 706)	TATGAACCACTAAGGCGTccaccttgttgGCCGGACCCAGCGCGACCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2279)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGGTCGCGCTGGGTCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3852)	1	CHD4_+_6716475.23-P1P2	CHD4	P1P2	na	na	na	na	na	na	na	na	na
GGGTCGGCTCCTCTGGGCAA (SEQ ID NO: 707)	TATGAACCACTAAGGCGTccaccttgttgGGGTCGGCTCCTCTGGGCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2280)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGCCCAGAGGAGCCGACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3853)	1	CHD4_-_6716355.23-P1P2	CHD4	P1P2	na	na	na	na	na	na	na	na	na
GTCCGGAGACGCCTGAGGAA (SEQ ID NO: 708)	TATGAACCACTAAGGCGTccaccttgttgGTCCGGAGACGCCTGAGGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2281)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCCTCAGGCGTCTCCGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3854)	1	CHD4_-_6716389.23-P1P2	CHD4	P1P2	na	na	na	na	na	na	na	na	na
GAGCCGGCCGAGGGTGGCGG (SEQ ID NO: 709)	TATGAACCACTAAGGCGTccaccttgttgGAGCCGGCCGAGGGTGGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2282)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCACCCTCGGCCGGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3855)	1	CHD5_-_6240357.23-P1P2	CHD5	P1P2	na	na	na	na	na	na	na	na	na
GAGGCGGCCGAGAGAAAGGG (SEQ ID NO: 710)	TATGAACCACTAAGGCGTccaccttgttgGAGGCGGCCGAGAGAAAGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2283)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCTTTCTCTCGGCCGCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3856)	1	CHD5_+_6240253.23-P1P2	CHD5	P1P2	na	na	na	na	na	na	na	na	na
GCGCAGCCGGCCGAGGGTGG (SEQ ID NO: 711)	TATGAACCACTAAGGCGTccaccttgttgGCGCAGCCGGCCGAGGGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2284)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACCCTCGGCCGGCTGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3857)	1	CHD5_-_6240354.23-P1P2	CHD5	P1P2	na	na	na	na	na	na	na	na	na
GCGGTCCGCCGAGGCCAGGG (SEQ ID NO: 712)	TATGAACCACTAAGGCGTccaccttgttgGCGGTCCGCCGAGGCCAGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2285)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCTGGCCTCGGCGGACCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3858)	1	CHD6_-_40246949.23-P1P2	CHD6	P1P2	na	na	na	na	na	na	na	na	na
GCTGCTGCAGGAGCCATTTT (SEQ ID NO: 713)	TATGAACCACTAAGGCGTccaccttgttgGCTGCTGCAGGAGCCATTTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2286)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAAAATGGCTCCTGCAGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3859)	1	CHD6_+_40246977.23-P1P2	CHD6	P1P2	na	na	na	na	na	na	na	na	na
GTGAATGGAGCCGGGATGAG (SEQ ID NO: 714)	TATGAACCACTAAGGCGTccaccttgttgGTGAATGGAGCCGGGATGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2287)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCATCCCGGCTCCATTCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3860)	1	CHD6_+_40246852.23-P1P2	CHD6	P1P2	na	na	na	na	na	na	na	na	na

GAAGGGCGGCCGGTAGAGTA (SEQ ID NO: 715)	TATGAACCACTAAGGCGTccaccttgttgGAAGGGCGGCCGGTAGAGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2288)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACTCTACCGGCCGCCCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3861)	2		CHEK1_-_125495934.23-P1P2	CHEK1	P1P2	na	na	na	na	na	na	na	na	na
GAGGATATAGGGAGCGGTAT (SEQ ID NO: 716)	TATGAACCACTAAGGCGTccaccttgttgGAGGATATAGGGAGCGGTATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2289)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATACCGCTCCCTATATCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3862)	2		CHEK1_+_125496272.23-P1P2	CHEK1	P1P2	na	na	na	na	na	na	na	na	na
GATGGGGCAGAAAACGCCTG (SEQ ID NO: 717)	TATGAACCACTAAGGCGTccaccttgttgGATGGGGCAGAAAACGCCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2290)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGGCGTTTTCTGCCCCATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3863)	2		CHEK1_+_125496255.23-P1P2	CHEK1	P1P2	na	na	na	na	na	na	na	na	na
GGAGAGTGTGCGGCTCCAGG (SEQ ID NO: 718)	TATGAACCACTAAGGCGTccaccttgttgGGAGAGTGTGCGGCTCCAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2291)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTGGAGCCGCACACTCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3864)	1		CHEK2_+_29137765.23-P1P2	CHEK2	P1P2	na	na	na	na	na	na	na	na	na
GGGCTCCAGGTGGGCTCACG (SEQ ID NO: 719)	TATGAACCACTAAGGCGTccaccttgttgGGGCTCCAGGTGGGCTCACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2292)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTGAGCCCACCTGGAGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3865)	1		CHEK2_+_29137755.23-P1P2	CHEK2	P1P2	na	na	na	na	na	na	na	na	na
GTCACGCGGTGAGTCATATG (SEQ ID NO: 720)	TATGAACCACTAAGGCGTccaccttgttgGTCACGCGGTGAGTCATATGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2293)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCATATGACTCACCGCGTGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3866)	1		CHEK2_+_29137741.23-P1P2	CHEK2	P1P2	na	na	na	na	na	na	na	na	na
GCCGCAATGTCTCTTGACAG (SEQ ID NO: 721)	TATGAACCACTAAGGCGTccaccttgttgGCCGCAATGTCTCTTGACAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2294)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGTCAAGAGACATTGCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3867)	1		CHFR_+_133464171.23-P1P2	CHFR	P1P2	na	na	na	na	na	na	na	na	na
GTGGAGACTCCCGACCCCAG (SEQ ID NO: 722)	TATGAACCACTAAGGCGTccaccttgttgGTGGAGACTCCCGACCCCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2295)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGGGGTCGGGAGTCTCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3868)	1		CHFR_-_133463971.23-P1P2	CHFR	P1P2	na	na	na	na	na	na	na	na	na
GTGTCTCTTGACAGCGGCGG (SEQ ID NO: 723)	TATGAACCACTAAGGCGTccaccttgttgGTGTCTCTTGACAGCGGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2296)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCCGCTGTCAAGAGACACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3869)	1		CHFR_+_133464165.23-P1P2	CHFR	P1P2	na	na	na	na	na	na	na	na	na
GAGGTTCGGAGCGGGAGCTC (SEQ ID NO: 724)	TATGAACCACTAAGGCGTccaccttgttgGAGGTTCGGAGCGGGAGCTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2297)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGCTCCCGCTCCGAACCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3870)	1		CHTF18_-_838671.23-P1P2	CHTF18	P1P2	na	na	na	na	na	na	na	na	na
GCTCGGGCTCGCGGACGGTA (SEQ ID NO: 725)	TATGAACCACTAAGGCGTccaccttgttgGCTCGGGCTCGCGGACGGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2298)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACCGTCCGCGAGCCCGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3871)	1		CHTF18_-_838687.23-P1P2	CHTF18	P1P2	na	na	na	na	na	na	na	na	na
GGGGCTCGCGGACGGTATGG (SEQ ID NO: 726)	TATGAACCACTAAGGCGTccaccttgttgGGGGCTCGCGGACGGTATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2299)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATACCGTCCGCGAGCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3872)	1		CHTF18_-_838690.23-P1P2	CHTF18	P1P2	na	na	na	na	na	na	na	na	na
GACGTGCTCGCTCACGGCGG (SEQ ID NO: 727)	TATGAACCACTAAGGCGTccaccttgttgGACGTGCTCGCTCACGGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2300)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCCGTGAGCGAGCACGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3873)	1		CLK2_+_155243267.23-P1P2	CLK2	P1P2	na	na	na	na	na	na	na	na	na
GCGGAGACCGAGCTGGGATC (SEQ ID NO: 728)	TATGAACCACTAAGGCGTccaccttgttgGCGGAGACCGAGCTGGGATCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2301)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGATCCCAGCTCGGTCTCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3874)	1		CLK2_+_155243222.23-P1P2	CLK2	P1P2	na	na	na	na	na	na	na	na	na
GCGTACGTGCTCGCTCACGG (SEQ ID NO: 729)	TATGAACCACTAAGGCGTccaccttgttgGCGTACGTGCTCGCTCACGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2302)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGTGAGCGAGCACGTACGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3875)	1		CLK2_+_155243270.23-P1P2	CLK2	P1P2	na	na	na	na	na	na	na	na	na
GCCGGGCGGGCGCCTCTCAC (SEQ ID NO: 730)	TATGAACCACTAAGGCGTccaccttgttgGCCGGGCGGGCGCCTCTCACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2303)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTGAGAGGCGCCCGCCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3876)	1		CLOCK_-_56412950.23-P1P2	CLOCK	P1P2	na	na	na	na	na	na	na	na	na
GCGCCGCCAAGGAAGCCAAC (SEQ ID NO: 731)	TATGAACCACTAAGGCGTccaccttgttgGCGCCGCCAAGGAAGCCAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2304)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTGGCTTCCTTGGCGGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3877)	1		CLOCK_-_56412906.23-P1P2	CLOCK	P1P2	na	na	na	na	na	na	na	na	na

GTCACCGGGGACGCGCTCG CGG (SEQ ID NO: 732)	TATGAACCACTAAGGCGTcaacctgtggGTCACCGGGCGAGCG CGCTCCGCGGTGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 2305)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACCCGCGAG 1 CLOCK_-5941295.2 3-P1P2 (SEQ ID NO: 3879)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCCCCTGCGCGTCCA CTA (SEQ ID NO: 733)	TATGAACCACTAAGGCGTcaacctgtggCCTCCCCTGCGC TCCACTAgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2306)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACTAGTGGA 2 CLSPN_-3926800.2 3-P1P2 GGCGCAGGGGGAGGCGGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3879)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CGCCGCTCTGCTCCGCGCC TGA (SEQ ID NO: 734)	TATGAACCACTAAGGCGTcaacctgtggCCTCCCCTGCGT CCGCGTGAgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2307)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACTCAGCCG 2 CLSPN_+3623519.23-P1P2 GAGAGCGAGCGGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3880)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGTCACGCGGGCGTCTA CCC (SEQ ID NO: 735)	TATGAACCACTAAGGCGTcaacctgtggCGGCTCACGCCGCGG TCTACCGgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2308)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACGGCTAGA 2 CLSPN_-3923582.2 3-P1P2 GCCGCGCTGACCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3881)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGCTCCGCGCACAGA TGG (SEQ ID NO: 736)	TATGAACCACTAAGGCGTcaacctgtggCGGGCTCCGCGCAC TCGCGGGgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2309)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACCCATCTTG 1 CREBBP_-3936096.23-P1P2 TCGCGGGACCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3882)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCTTCGGAGCGCGATCT ACT (SEQ ID NO: 737)	TATGAACCACTAAGGCGTcaacctgtggGGCTTCCGAGCGCGG ATCTACTgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2310)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACAGTAGATG 1 CREBBP_+3936647.23-P1P2 GCGCTCGAAGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3883)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGCGCTCGAAGCGCCC GGT (SEQ ID NO: 738)	TATGAACCACTAAGGCGTcaacctgtggCTCGCGCTCGAAG CCCCGGTgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2311)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACACCGGGG 1 CREBBP_-3936877.23-P1P2 CTTCGAGCGCGGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3884)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAATCTTGCCTGTCTCAG CAG (SEQ ID NO: 739)	TATGAACCACTAAGGCGTcaacctgtggGGAAGTCTTGCCTGT CTGCAGgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2312)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACCTGGTGA 1 CXorf57_-109885208 23-P1P2 CCAGGCAAAGCTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3885)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAGTTTGCCCCGTCCAA AG (SEQ ID NO: 740)	TATGAACCACTAAGGCGTcaacctgtggGTAGTTTGCCCCGT TGGGGCAAACTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 2313)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACCTTTGGAC 1 CXorf57_+109885251.23-P1P2 TGGGGCAAACTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3886)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTGCAGCCTCCGCTAGG GGG (SEQ ID NO: 741)	TATGAACCACTAAGGCGTcaacctgtggCTTGCAGCCGTCGC TACGAGgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2314)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACCCCCGTA 1 CXorf57_-106955187 23-P1P2 GACGAGCTGCAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3887)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGCGCAGTCGTGCACCC CAA (SEQ ID NO: 742)	TATGAACCACTAAGGCGTcaacctgtggCGCTGCCACTCGCGC ACCCCAAgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2315)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTTAAACTTGGGGT 1 DCLRE1A_-115613633 23-P1P2 GCCGAAGTCGGCAGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3888)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGCGGTGAGGGTGCCT TGA (SEQ ID NO: 743)	TATGAACCACTAAGGCGTcaacctgtggGGGGCGGTGAGGG TCCCTGAgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2316)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACTCCAAGGA 1 DCLRE1A_+115613895 23-P1P2 CCCCTCACCCGCCCCCAAGAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3889)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTTCCCGCCAACGACCC TAA (SEQ ID NO: 744)	TATGAACCACTAAGGCGTcaacctgtggCGTTCCCCGCCAAG ACCCCTAAgttaagagctaagtgGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2317)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACTTAGGGTC 1 DCLRE1A_+115614111 23-P1P2 TTGGCCGGGAAACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3890)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACAGGGCCACGACGACCC TAA (SEQ ID NO: 745)	TATGAACCACTAAGGCGTcaacctgtggGACGAGGCCACGAG AGCCCTAAgttaagagctaagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2318)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACGTTAGGGTC 1 DCLRE1B_+114448033 23-P1P2 TCTGGGCCCTGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3891)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGCTGGAGCCCAAA TCG (SEQ ID NO: 746)	TATGAACCACTAAGGCGTcaacctgtggGCCGCTGGAGCCC CAAATCGgttaagagctaagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2319)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACGGATTTGG 1 DCLRE1B_+114448034 23-P1P2 GCTCCACGCGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3892)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTGGACGCCAAATCG CGG (SEQ ID NO: 747)	TATGAACCACTAAGGCGTcaacctgtggCGGTCGACGCCAA ATCGCGGgttaagagctaagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2320)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACCCGCGAT 1 DCLRE1B_+114448501 23-P1P2 TGTGGCTCCCAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3893)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGCGCCGATCCCGAGA GTC (SEQ ID NO: 748)	TATGAACCACTAAGGCGTcaacctgtggCGCGCGCCGATCC CAAGAGTCgttaagagctaagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2321)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACGACTCTCTG 1 DCLRE1C_-149960322 3-P1P2 GGATCGCGCGCCGATCCCGAGAGTCTTAGTGGTTCAT A (SEQ ID NO: 3894)	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGGAAGTGGCGGCGCGGTCA (SEQ ID NO: 749)	TATGAACCACTAAGGCGTccaccttgttgGGGAAGTGGCGGCGCGGTCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2322)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGACCGCGCCGCCACTTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3895)	1		DCLRE1C_+_14996065.23-P1P2	DCLRE1C	P1P2	na	na	na	na	na	na	na	na	na
GGGGTCCCGGACTCTGGGAT (SEQ ID NO: 750)	TATGAACCACTAAGGCGTccaccttgttgGGGGTCCCGGACTCTGGGATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2323)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATCCCAGAGTCCGGGACCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3896)	1		DCLRE1C_+_14996016.23-P1P2	DCLRE1C	P1P2	na	na	na	na	na	na	na	na	na
GGCCGGGGCGCGGGGCATCA (SEQ ID NO: 751)	TATGAACCACTAAGGCGTccaccttgttgGGCCGGGGCGCGGGGCATCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2324)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGATGCCCCGCGCCCCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3897)	1		DDB1_-_61100604.23-P1	DDB1	P1	na	na	na	na	na	na	na	na	na
GGGCCTGCAACGTGTCTCTG (SEQ ID NO: 752)	TATGAACCACTAAGGCGTccaccttgttgGGGCCTGCAACGTGTCTCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2325)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGAGACACGTTGCAGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3898)	1		DDB1_+_61100555.23-P1	DDB1	P1	na	na	na	na	na	na	na	na	na
GGTCGCGCTCGAGTCCCGAC (SEQ ID NO: 753)	TATGAACCACTAAGGCGTccaccttgttgGGTCGCGCTCGAGTCCCGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2326)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCGGGACTCGAGCGCGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3899)	1		DDB1_+_61100457.23-P1	DDB1	P1	na	na	na	na	na	na	na	na	na
GGAAAAGTCCTCAATACGGT (SEQ ID NO: 754)	TATGAACCACTAAGGCGTccaccttgttgGGAAAAGTCCTCAATACGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2327)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCGTATTGAGGACTTTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3900)	1		DDB1_-_61081136.23-P2	DDB1	P2	na	na	na	na	na	na	na	na	na
GGAACGAAAAGTCCTCAATA (SEQ ID NO: 755)	TATGAACCACTAAGGCGTccaccttgttgGGAACGAAAAGTCCTCAATAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2328)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTATTGAGGACTTTTCGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3901)	1		DDB1_-_61081132.23-P2	DDB1	P2	na	na	na	na	na	na	na	na	na
GTTGGGCTCAACATTGAGAC (SEQ ID NO: 756)	TATGAACCACTAAGGCGTccaccttgttgGTTGGGCTCAACATTGAGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2329)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCTCAATGTTGAGCCCAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3902)	1		DDB1_+_61081332.23-P2	DDB1	P2	na	na	na	na	na	na	na	na	na
GCTTCACACGGAGGACGCGA (SEQ ID NO: 757)	TATGAACCACTAAGGCGTccaccttgttgGCTTCACACGGAGGACGCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2330)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCGTCCTCCGTGTGAAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3903)	1		DDB2_-_47236690.23-P1P2	DDB2	P1P2	na	na	na	na	na	na	na	na	na
GGAGGATTGAGAGGCCTCTG (SEQ ID NO: 758)	TATGAACCACTAAGGCGTccaccttgttgGGAGGATTGAGAGGCCTCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2331)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGAGGCCTCTCAATCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3904)	1		DDB2_+_47236615.23-P1P2	DDB2	P1P2	na	na	na	na	na	na	na	na	na
GTCTATGCGAAGATCATGGA (SEQ ID NO: 759)	TATGAACCACTAAGGCGTccaccttgttgGTCTATGCGAAGATCATGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2332)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCATGATCTTCGCATAGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3905)	1		DDB2_+_47236636.23-P1P2	DDB2	P1P2	na	na	na	na	na	na	na	na	na
GAGGCGAAGCCGCGGAGGAC (SEQ ID NO: 760)	TATGAACCACTAAGGCGTccaccttgttgGAGGCGAAGCCGCGGAGGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2333)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCCTCCGCGGCTTCGCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3906)	1		DDX1_-_15732050.23-P1P2	DDX1	P1P2	na	na	na	na	na	na	na	na	na
GCGAGCAGGCGAAGCCGCGG (SEQ ID NO: 761)	TATGAACCACTAAGGCGTccaccttgttgGCGAGCAGGCGAAGCCGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2334)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCGGCTTCGCCTGCTCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3907)	1		DDX1_-_15732045.23-P1P2	DDX1	P1P2	na	na	na	na	na	na	na	na	na
GGAGCGAGCAGGCGAAGCCG (SEQ ID NO: 762)	TATGAACCACTAAGGCGTccaccttgttgGGAGCGAGCAGGCGAAGCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2335)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGCTTCGCCTGCTCGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3908)	1		DDX1_-_15732042.23-P1P2	DDX1	P1P2	na	na	na	na	na	na	na	na	na
GACGCCACTGAGTTAGAAAC (SEQ ID NO: 763)	TATGAACCACTAAGGCGTccaccttgttgGACGCCACTGAGTTAGAAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2336)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTTCTAACTCAGTGGCGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3909)	1		DDX11_+_31226826.23-P1P2	DDX11	P1P2	na	na	na	na	na	na	na	na	na
GCCACTGAGTTAGAAACTGG (SEQ ID NO: 764)	TATGAACCACTAAGGCGTccaccttgttgGCCACTGAGTTAGAAACTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2337)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCAGTTTCTAACTCAGTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3910)	1		DDX11_+_31226823.23-P1P2	DDX11	P1P2	na	na	na	na	na	na	na	na	na
GGCAGCAGCGAGAATCTACA (SEQ ID NO: 765)	TATGAACCACTAAGGCGTccaccttgttgGGCAGCAGCGAGAATCTACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2338)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTAGATTCTCGCTGCTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3911)	1		DDX11_+_31226729.23-P1P2	DDX11	P1P2	na	na	na	na	na	na	na	na	na

GATCTGCCTAGCTCCCACTA (SEQ ID NO: 766)	TATGAACCACTAAGGCGTccaccttgttgGATCTGCCTAGCTCCCACTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2339)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGTGGGAGCTAGGCAGATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3912)	1		DDX3X_-_41192858.23-P1P2	DDX3X	P1P2	na	na	na	na	na	na	na	na	na
GCCCTCTCACGGGAGAACTG (SEQ ID NO: 767)	TATGAACCACTAAGGCGTccaccttgttgGCCCTCTCACGGGAGAACTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2340)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGTTCTCCCGTGAGAGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3913)	1		DDX3X_+_41193418.23-P1P2	DDX3X	P1P2	na	na	na	na	na	na	na	na	na
GTCCCGTGAGAGGGCCTTCG (SEQ ID NO: 768)	TATGAACCACTAAGGCGTccaccttgttgGTCCCGTGAGAGGGCCTTCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2341)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGAAGGCCCTCTCACGGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3914)	1		DDX3X_-_41193448.23-P1P2	DDX3X	P1P2	na	na	na	na	na	na	na	na	na
GAAGCTGCACTACTAGAGAC (SEQ ID NO: 769)	TATGAACCACTAAGGCGTccaccttgttgGAAGCTGCACTACTAGAGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2342)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCTCTAGTAGTGCAGCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3915)	1		DDX5_-_62502390.23-P1P2	DDX5	P1P2	na	na	na	na	na	na	na	na	na
GAGACCGGTAGAAATGAATG (SEQ ID NO: 770)	TATGAACCACTAAGGCGTccaccttgttgGAGACCGGTAGAAATGAATGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2343)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCATTCATTTCTACCGGTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3916)	1		DDX5_-_62502405.23-P1P2	DDX5	P1P2	na	na	na	na	na	na	na	na	na
GTGACGGCGAAGCCTTGCGG (SEQ ID NO: 771)	TATGAACCACTAAGGCGTccaccttgttgGTGACGGCGAAGCCTTGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2344)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCAAGGCTTCGCCGTCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3917)	1		DDX5_-_62502333.23-P1P2	DDX5	P1P2	na	na	na	na	na	na	na	na	na
GCGTAGCTAGCTGCAAGCAG (SEQ ID NO: 772)	TATGAACCACTAAGGCGTccaccttgttgGCGTAGCTAGCTGCAAGCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2345)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGCTTGCAGCTAGCTACGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3918)	1		DDX58_+_32526169.23-P1P2	DDX58	P1P2	na	na	na	na	na	na	na	na	na
GGCTAGCTGCAAGCAGAGGC (SEQ ID NO: 773)	TATGAACCACTAAGGCGTccaccttgttgGGCTAGCTGCAAGCAGAGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2346)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCTCTGCTTGCAGCTAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3919)	1		DDX58_+_32526165.23-P1P2	DDX58	P1P2	na	na	na	na	na	na	na	na	na
GTAGTGAGGCACAGCCTGCG (SEQ ID NO: 774)	TATGAACCACTAAGGCGTccaccttgttgGTAGTGAGGCACAGCCTGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2347)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCAGGCTGTGCCTCACTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3920)	1		DDX58_+_32526193.23-P1P2	DDX58	P1P2	na	na	na	na	na	na	na	na	na
GACCGTGTGGGCTGTAGTAG (SEQ ID NO: 775)	TATGAACCACTAAGGCGTccaccttgttgGACCGTGTGGGCTGTAGTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2348)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTACTACAGCCCACACGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3921)	2		DHX15_+_24586116.23-P1P2	DHX15	P1P2	na	na	na	na	na	na	na	na	na
GGGTTCCGGCCGCACCGTGT (SEQ ID NO: 776)	TATGAACCACTAAGGCGTccaccttgttgGGGTTCCGGCCGCACCGTGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2349)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACACGGTGCGGCCGGAACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3922)	2		DHX15_+_24586128.23-P1P2	DHX15	P1P2	na	na	na	na	na	na	na	na	na
GTGTGGGCTGTAGTAGCGGG (SEQ ID NO: 777)	TATGAACCACTAAGGCGTccaccttgttgGTGTGGGCTGTAGTAGCGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2350)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCTACTACAGCCCACACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3923)	2		DHX15_+_24586112.23-P1P2	DHX15	P1P2	na	na	na	na	na	na	na	na	na
GAGCGGTTGCTGGTTCTGAC (SEQ ID NO: 778)	TATGAACCACTAAGGCGTccaccttgttgGAGCGGTTGCTGGTTCTGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2351)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCAGAACCAGCAACCGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3924)	1		DHX36_+_154042223.23-P1P2	DHX36	P1P2	na	na	na	na	na	na	na	na	na
GGCCCGGACGCCATTTCCAG (SEQ ID NO: 779)	TATGAACCACTAAGGCGTccaccttgttgGGCCCGGACGCCATTTCCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2352)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGGAAATGGCGTCCGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3925)	1		DHX36_+_154042240.23-P1P2	DHX36	P1P2	na	na	na	na	na	na	na	na	na
GGGCATCCCGGGCACCTGAA (SEQ ID NO: 780)	TATGAACCACTAAGGCGTccaccttgttgGGGCATCCCGGGCACCTGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2353)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCAGGTGCCCGGGATGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3926)	1		DHX36_+_154042029.23-P1P2	DHX36	P1P2	na	na	na	na	na	na	na	na	na
GCGCTCAGAGCTACTCCGGC (SEQ ID NO: 781)	TATGAACCACTAAGGCGTccaccttgttgGCGCTCAGAGCTACTCCGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2354)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCGGAGTAGCTCTGAGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3927)	2		DHX8_+_41561345.23-P1P2	DHX8	P1P2	na	na	na	na	na	na	na	na	na
GGCGGGAGCCTTAATCGGGT (SEQ ID NO: 782)	TATGAACCACTAAGGCGTccaccttgttgGGCGGGAGCCTTAATCGGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2355)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCCGATTAAGGCTCCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3928)	2		DHX8_-_41561447.23-P1P2	DHX8	P1P2	na	na	na	na	na	na	na	na	na

GGCTCTGAGCGCGGCGGCT (SEQ ID NO: 783)	TATGAAGCACTAAGGCGTcaacttgtgGGCTGTGAGGCGCCG3GCTGTCgttaagagtaagtgtCGTAAAGTCTGTGCTGGTGA (SEQ ID NO: 2356)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCACACGCC	2	DHX8_-_41561378.2 3-P1P2	DHX8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAACGCACAAGCAACTC (SEQ ID NO: 784)	TATGAAGCACTAAGGCGTcaacttgtgGAAACGCACAAGCAA ACTGTGCAgttaagagtaagtgtCGTAAAGTCTGTGTGTGA (SEQ ID NO: 2357)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTGCGAGTT		DHX9_+_182803499.23-P1P2	DHX9	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CAGAAGGCCAGACATCAC (SEQ ID NO: 785)	TATGAAGCACTAAGGCGTcaacttgtgAGAAGCCCACAG TGTGTGCAgttaagagtaagtgtCGTAAAGTCTGTGTGTGA (SEQ ID NO: 2358)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTGTGTGCAG		DHX9_-_192858557.23-P1P2	DHX9	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGAAATGGCGTCCGAA (SEQ ID NO: 786)	TATGAAGCACTAAGGCGTcaacttgtgCGGCGAAATGGCT ACGCATTTCGACCCAAGAGTTGGACTGCTTAGTGGTTCATA (SEQ ID NO: 2359)	TCACCAGCAGAGTTTAGCGCAGGTTAGCTCTTAAACCCGTCGC		DHX9_+_182809479.23-P1P2	DHX9	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATGGAGTAACGTGACA (SEQ ID NO: 787)	TATGAACCACTAAGGCGTcaacttgtgATGGAGTAACGT CACAAGCGgttaagagtaagtgtCGTAAAGTCTGTGTGTGA (SEQ ID NO: 2360)	TCACCAGCAGAGTTTAGCGCAGGTTAGCTCTTAAACGCTGTC AGGTTACTCCATCGAAAGAGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3633)		DICER1_+_9582370723-P1P2	DICER1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAACTGCGGCGCCAACG (SEQ ID NO: 788)	TATGAAGCACTAAGGCGTcaacttgtgGCAACTGCGGCGCCC AAAGCGCAgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2361)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTGGCGTT GGGCGCGAAGTTGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3634)		DICER1_+_9582405523-P1P2	DICER1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAACCTGACAGCGGGG (SEQ ID NO: 789)	TATGAAGCACTAAGGCGTcaacttgtgGTAACCTGACAGCA GGGCGACCgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2362)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACGTCCCC GCTGTCAAGTTACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3635)		DICER1_+_9582395523-P1P2	DICER1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAAGCTGCCGGGACCCAG (SEQ ID NO: 790)	TATGAACCACTAAGGCGTcaacttgtgGCAAGCTGCCGGGAC CCCAGGTTgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2363)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCACGTGG GTCCGGGCAGCGTCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3636)		DNAP1_-_44679264.23-P1P2	DNAP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGGGCTTTGCGGGGA (SEQ ID NO: 791)	TATGAACCACTAAGGCGTcaacttgtgGCGGGGCTTTGCG GGGACGCGgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2364)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCCGTCGC CGCAAAGGCCGCGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3637)		DNAP1_-_44679185.23-P1P2	DNAP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGAGCCACCGCGAGGT (SEQ ID NO: 792)	TATGAAGCACTAAGGCGTcaacttgtgCGGGGAGCCACCGCG AAGGTCAAgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2365)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTTGACCTC CGGGTGGCTCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3638)		DNAP1_+_44679387.23-P1P2	DNAP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGTTTGGCGCCAAAGCC (SEQ ID NO: 793)	TATGAACCACTAAGGCGTcaacttgtgCGCGTTTGGCGCGCA AACCCGTCgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2366)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACGACGGGTTT GGGCGCCAAAAGGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3639)		DMC1_+_38966174.23-P1P2	DMC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTAACTCGGAGGCCTAGA (SEQ ID NO: 794)	TATGAACCACTAAGGCGTcaacttgtgGCTAACTCGGAGGCC TAGAAACAgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2367)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACTGTTCTAG GCCCGAGTTAGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3640)		DMC1_-_38986148.23-P1P2	DMC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCTGAACAGGAATTCTAAC (SEQ ID NO: 795)	TATGAAGCACTAAGGCGTcaacttgtgCGGCTGAACAGGAAT TCTAACTTgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2368)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACGAGTTA ATCGCTGTGAGCCCAAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3641)		DMC1_-_38968138.23-P1P2	DMC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGGGCCCGGTCAGT (SEQ ID NO: 796)	TATGAACCACTAAGGCGTcaacttgtgGCCCGGGCCCGGGT CAGTGTCTgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2369)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACGCAGAGTG AGCCGGGCCCGGGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 3642)		DNA2_-_70231468.23-P1P2	DNA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGAGCTGTGAGCGGAG (SEQ ID NO: 797)	TATGAACCACTAAGGCGTcaacttgtgGGGAGCTGTGAGC GGAGAGGTgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2370)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCGTCTCCGG CTCACAGCTCCCCGAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3643)		DNA2_+_70231531.23-P1P2	DNA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGTCCCCGTGTCAGTCTG (SEQ ID NO: 798)	TATGAAGCACTAAGGCGTcaacttgtgTGTCCCCGTGTCAGT CTGTGAGCgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2371)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCCTCAGAC TGAGCCGGGCCCCCAAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3644)		DNA2_-_70231471.23-P1P2	DNA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTACGCGCGCGGCATCT (SEQ ID NO: 799)	TATGAACCACTAAGGCGTcaacttgtgGTACGCGCGCGGCA TCTCTCGGgttaagagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2372)	TCACCAGCAGAGCTTTAGCGCAGGTTAGCTCTTAAACCCGAGAT GGCGCGCGCGTACCCAACAAGGTCGACGCCTTAGTCGGTTCAT A (SEQ ID NO: 3645)		DNMT1_-_10305882.23-P1P2	DNMT1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTGGGAGGCTTCAGCAG AG3 (SEQ ID NO: 800)	TATGAAGCACTAAG3GCTcaacttgtgG TCGGAG3CTTCA G3AAG3CCTCCCGACCAACAAGGTG3AG3GCCTTAGTCGTTCATA (SEQ ID NO: 3946)		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTGCT	DNMT1_-_10305697.2 3-P1P2	DNMT1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGCGGTGGACGGGCCAG TGT (SEQ ID NO: 801)	TATGAACCACTAAG3GCTcaacttgtgG TCGGAG3CTTCA GCGTCCCCTGGCCAGCAAGGTCGACG3CCTTAGTGGTTCATA (SEQ ID NO: 3947)		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTGCT	DNMT1_-_10305644.2 3-P1P2	DNMT1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
CCACTCCTCGAGGGGCTC CAG (SEQ ID NO: 802)	TATGAACCACTAAG3GCTcaacttgtgG TCGGAG3CTTCA GCTCCGAGGAGTGTCCAAGAAGTCTGCTGGTGA (SEQ ID NO: 3948)		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCGAG3C	DNMT3A_+_25474534.2 3-ENST00000402867.1	DNMT3A	ENST0000040286 71	na	na	na	na	na	na	na	na	na	na	na	na
GCCAACGGCCCCATCCCT CGA (SEQ ID NO: 803)	TATGAACCACTAAG3GCTcaacttgtgG TCCATCGAGTGTGGCAAGAAGTCTGCTGGTGA (SEQ ID NO: 3949)		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCGAG3C	DNMT3A_+_25474843.2 3-ENST00000402867.1	DNMT3A	ENST0000040286 71	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCCGCTGGAGGGCCT GGT (SEQ ID NO: 804)	TATGAACCACTAAG3GCTcaacttgtgG TCCCTCGAGTGTGGCAAGAAGTCTGCTGGTGA (SEQ ID NO: 3950)		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCGAG3C	DNMT3A_-_25475282.2 3-ENST00000402867.1	DNMT3A	ENST0000040286 71	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAAGCCCTCCGAAGCAGC TGG (SEQ ID NO: 805)	TATGAACCACTAAG3GCTcaacttgtgG TCAACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTG		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTG	DNMT3B_+_31350536.2 3-P1	DNMT3B	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCACTCCCCGCTCGCCCC GTC (SEQ ID NO: 806)	TATGAAGCACTAAG3GCTcaacttgtgG TCCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTG		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTG	DNMT3B_-_31365213.2 3-P1	DNMT3B	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACCGGTCCCGTTCCGTC GGG (SEQ ID NO: 807)	TATGAACCACTAAG3GCTcaacttgtgG TCCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTG		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTG	DNMT3B_+_31350282.2 3-P1	DNMT3B	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCACCCCCGGAGATCAG AGG (SEQ ID NO: 808)	TATGAACCACTAAG3GCTcaacttgtgG TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTG		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCCTGTCTG	DNMT3B_-_31368274.2 3-P2	DNMT3B	P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGATAGCCTCCAGGAT TGG (SEQ ID NO: 809)	TATGAAGCACTAAG3GCTcaacttgtgG TCAGCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACCCAA1GCT		TCAGCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACCCAA1GCT	DNMT3B_+_31369231.23 3-P2	DNMT3B	P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCGACACCCCGGAGAT AGG (SEQ ID NO: 810)	TATGAACCACTAAG3GCTcaacttgtgG TCCACCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACCTGATCTC		TCACCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACCTGATCTC	DNMT3B_-_31368271.2 3-P2	DNMT3B	P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGCTCGTGGTGGGATC CAT (SEQ ID NO: 811)	TATGAAGCACTAAG3GCTcaacttgtgG TACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACATGGATCC		TCACCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACATGGATCC	DNTT_+_98364251.23-P1P2	DNTT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCACCAGATGGGCCCAGC CAG (SEQ ID NO: 812)	TATGAAGCACTAAG3GCTcaacttgtgG CACCAGAGTGGGCCCAGC		TCACCAGCAGAGACTTTACG3CAG3CTTAG3CTCTTAAACCTGGCTG	DNTT_-_98064235.2 3-P1P2	DNTT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTCTTCCGAGGGCTCAA CTC (SEQ ID NO: 813)	TATGAACCACTAAG3GCTcaacttgtgG TTCTTCGGAGGGCTCAAGCGG		TCACCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACACTTGAGCC	DNTT_+_96364273.23-P1P2	DNTT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCATGTTTGCGACTCCTTG GTG (SEQ ID NO: 814)	TATGAACCACTAAG3GCTcaacttgtgG CATGTTGCGACTCCTTGGTGA		TCACCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACCACCAG3G	DPF2_-_85101522.2 3-P1	DPF2	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGGCGCGTCCGGGACT GTG (SEQ ID NO: 815)	TATGAAGCACTAAG3GCTcaacttgtgG CGTGCGGCTCCGGGACTGGTG		TCAGCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACCACGCTCCC	DPF2_-_85101329.2 3-P1	DPF2	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
CGCCTTCTCGGCCCGAGCG CAG (SEQ ID NO: 816)	TATGAACCACTAAG3GCTcaacttgtgG TCGCCTTCTCGGCCCGAGCG		TCACCAGCAGACTTTACG3CAG3CTTAG3CTCTTAAACCTTGCGTC	DPF2_-_85101345.2 3-P1	DPF2	P1	na	na	na	na	na	na	na	na	na	na	na	na	na

GCAGGAATGGTCAGTAGTT TAC (SEQ ID NO: 817)	TATGAAACCACTAAGGCGTcaacctgtggCAGGATGGTCAGT ACTTTTACgattaagatcaagcgtgCGTAAAAGTCTGCTGCTTGA (SEQ ID NO: 2390)	TCAGCAGCAGACTTTACGCAGCTTAGCTATCGTCTTAAACTCTAAACTA CTGCCATCGTCGCCAACGAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 3963)	DPF2_-_85108628.2 3-P2	DPF2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGTAGCTGACAATAAT CA (SEQ ID NO: 818)	TATGAAACCACTAAGGCGTcaacctgtgCAGTAGCTGACAA TAATCAgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2391)	TCAGCAGCAGACTTTACGCAGCTTAGCTATCGTCTTAAACTGATTATT GTCAGGTTACTGCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3964)	DPF2_+_8617123-P2	DPF2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTCACAGCAAAACTCG CCC (SEQ ID NO: 819)	TATGAAACCACTAAGGCGTcaacctgtgGGTCACAGCAAA TTTGCTGTGAACCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 2392)	TCAACCAGCAGACTTTACGCAGCTTAGCTGTTAAACGGGCGCAG TTTGCTGTGAACCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3965)	DPF2_+_86198993.23-P2	DPF2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAAGCGCAGCGCTACTAC CGG (SEQ ID NO: 820)	TATGAAACCACTAAGGCGTcaacctgtgCAAAGCCACGCGTA TAGCCTGGACTTTCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 2393)	TCAGCAGCAGACTTTACGCAGCTTAGCTGTTAAACGTCGGTTTC TAGCCTGGACTTTCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3966)	DROSHA_-_31532158.2 3-P1P2	DROSHA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GACCGACAGGTACCAAGA AAG (SEQ ID NO: 821)	TATGAAACCACTAAGGCGTcaacctgtgGACCGACAGGTA AACAACACgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2394)	TCAACCAGCAGACTTTACGCAGCTTAGCTGTTAAACTCGGTTTC GTACCTGCGGTTCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3967)	DROSHA_-_31532170.2 3-P1P2	DROSHA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCCGGGAACACGGAAAC CGA (SEQ ID NO: 822)	TATGAAACCACTAAGGCGTcaacctgtgGCTCCGGGAACACAGG AAACCGAgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2395)	TCAACCAGCAGACTTTACGCAGCTTAGCTGTTAAACTCGGTTTCC CTGTTCCGGAGCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3968)	DROSHA_-_31532264.2 3-P1P2	DROSHA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAACAGGGCTTGACAGC GCG (SEQ ID NO: 823)	TATGAAACCACTAAGGCGTcaacctgtgGAACAGGGCTTGA CAGCGCGAgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2396)	TCAGCAGCAGACTTTACGCAGCTTAGCTATCGTCTTAAACTCGCGCTG TCAACGCCTGTTCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3969)	DTL_+_21223098975.23-P1P2	DTL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACAAACCGCGGGGGTG ACC (SEQ ID NO: 824)	TATGAAACCACTAAGGCGTcaacctgtgGGACAAACCGCGGG GCGTGAACgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2397)	TCAACCAGCAGACTTTACGCAGCTTAGCTGTTAAACGGGTCACCC CCCGCGTTTGTCCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3970)	DTL_+_21223098907.23-P1P2	DTL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTACAGTGGCGGGAGT TGG (SEQ ID NO: 825)	TATGAAACCACTAAGGCGTcaacctgtgGTTACAGTGGCGG TGGGTTGGgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2398)	TCAACCAGCAGACTTTACGCAGCTTAGCTATCGTCTTAAACCCAACTCC CGCCACTGTTAACCACAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3971)	DTL_-_21220991282.3-P1P2	DTL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGAGAACACCGGGGGG AAG (SEQ ID NO: 826)	TATGAAACCACTAAGGCGTcaacctgtgGAGGAGAACACCGG GGGGAAGgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2399)	TCAACCAGCAGACTTTACGCAGCTTAGCTATCGTCTTAAACCTTCCCCCC CGGTGGTCCGTCCAACCAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3972)	DUT_+_46824405.23-P1P2	DUT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGCGAGCGAAGGAGAC CAC (SEQ ID NO: 827)	TATGAAACCACTAAGGCGTcaacctgtgGAGGGCGAGCGAA GGAGACCACgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2400)	TCAACCAGCAGACTTTACGCAGCTTAGCTATCGTCTTAAAGGTGTGTGT CCTCGGCTCGCTGCGCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3973)	DUT_+_46824414.23-P1P2	DUT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGAAGCGAGAGAGACGAC CGG (SEQ ID NO: 828)	TATGAAACCACTAAGGCGTcaacctgtgGCGAAGCGAGAGAGA CCACCGCgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2401)	TCAACCAGCAGACTTTACGCAGCTTAGCTCTGTTAAACCCCGCGTGG TCTCCTCGCTCGCGCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3974)	DUT_+_46824411.23-P1P2	DUT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGCCCGCGCTGACAGC CGG (SEQ ID NO: 829)	TATGAAACCACTAAGGCGTcaacctgtgCGgctcacggtgctgacaggtc gatttaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2402)	TCAACCAGCAGACTTTACGCAGCTTAGCTATCGTCTTAAACCCCGCCTG TCAACCCGGGTGCCGGCCAACAAGGTTGCAGCGCCTTAGTGGTTCAT A (SEQ ID NO: 3975)	E2F1_-_32274108.23-P1P2	E2F1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGCTGCTGCAAAG TCC (SEQ ID NO: 830)	TATGAAACCACTAAGGCGTcaacctgtgGCCGCTGCTGC AAAAGTCCgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2403)	TCAACCAGCAGACTTTACGCAGCTTAGCTGTTAAACGGGAGCTTTG CAAGCCAGTCGCCCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3976)	E2F1_-_32274182.23-P1P2	E2F1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGCGGCAGGGCGCTCGG CGA (SEQ ID NO: 831)	TATGAAACCACTAAGGCGTcaacctgtgGTGGCGGCAGGGC TCGGCGAgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2404)	TCAACCAGCAGACTTTACGCAGCTTAGCTGTTAAACTCCGCCGA TGCCGTCGCGGCCAACCAAGGTTGCAGCGCCTTAGTGGTTCAT A (SEQ ID NO: 3977)	E2F1_-_32274145.23-P1P2	E2F1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGAAACGTCTTTGGAA GG (SEQ ID NO: 832)	TATGAAACCACTAAGGCGTcaacctgtgGCTGAAACGTCTTT GGAAGGgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2405)	TCAACCAGCAGACTTTACGCAGCTTAGCTATCGTCTTAAACGTTTCCAA AGACGTTTCAGCCGAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3978)	EED_-_85965876.23-P1P2	EED	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAGCGGGGTTGGACAT CGA (SEQ ID NO: 833)	TATGAAACCACTAAGGCGTcaacctgtgGGCAGCGGGGTTGG ACATCGAgattaagatcaagctgtgCGTAAAAGTCTGCTGGTGA (SEQ ID NO: 2406)	TCAACCAGCAGACTTTACGCAGCTTAGCTCTGTTAAACGTCGATGTC CGAGCCCGTCTGCCCAACAAGGTTGCAGCGCCTTAGTGGTTCATA (SEQ ID NO: 3979)	EED_-_85965883.23-P1P2	EED	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na

GGTCTTTGGAAGGAGGAAGG (SEQ ID NO: 834)	TATGAACCACTAAGGCGTccaccttgttgGGTCTTTGGAAGGAGGAAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2407)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTTCCTCCTTCCAAAGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3980)	1	EED_-_85955883.23-P1P2	EED	P1P2	na	na	na	na	na	na	na	na	na
GATCTCTGGTGGCTCTGCTA (SEQ ID NO: 835)	TATGAACCACTAAGGCGTccaccttgttgGATCTCTGGTGGCTCTGCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2408)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGCAGAGCCACCAGAGATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3981)	1	ELP3_-_27950720.23-P1P2	ELP3	P1P2	na	na	na	na	na	na	na	na	na
GCTACGGCGGCGCAGAAATG (SEQ ID NO: 836)	TATGAACCACTAAGGCGTccaccttgttgGCTACGGCGGCGCAGAAATGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2409)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCATTTCTGCGCCGCCGTAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3982)	1	ELP3_-_27950736.23-P1P2	ELP3	P1P2	na	na	na	na	na	na	na	na	na
GTCTGGTGGCTCTGCTACGG (SEQ ID NO: 837)	TATGAACCACTAAGGCGTccaccttgttgGTCTGGTGGCTCTGCTACGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2410)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGTAGCAGAGCCACCAGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3983)	1	ELP3_-_27950723.23-P1P2	ELP3	P1P2	na	na	na	na	na	na	na	na	na
GCCACCAAGATTCTGCAGAG (SEQ ID NO: 838)	TATGAACCACTAAGGCGTccaccttgttgGCCACCAAGATTCTGCAGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2411)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCTGCAGAATCTTGGTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3984)	1	EME1_+_48450692.23-P1P2	EME1	P1P2	na	na	na	na	na	na	na	na	na
GGAGTGGCGGGAGAAGTTGC (SEQ ID NO: 839)	TATGAACCACTAAGGCGTccaccttgttgGGAGTGGCGGGAGAAGTTGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2412)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCAACTTCTCCCGCCACTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3985)	1	EME1_-_48450638.23-P1P2	EME1	P1P2	na	na	na	na	na	na	na	na	na
GTGCGTGGCAGTTGAAAGAG (SEQ ID NO: 840)	TATGAACCACTAAGGCGTccaccttgttgGTGCGTGGCAGTTGAAAGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2413)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCTTTCAACTGCCACGCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3986)	1	EME1_-_48450622.23-P1P2	EME1	P1P2	na	na	na	na	na	na	na	na	na
GAGCGAGGAAGAGGTCGGTC (SEQ ID NO: 841)	TATGAACCACTAAGGCGTccaccttgttgGAGCGAGGAAGAGGTCGGTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2414)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGACCGACCTCTTCCTCGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3987)	1	EME2_-_1823225.23-P1P2	EME2	P1P2	na	na	na	na	na	na	na	na	na
GCCGGAAGCGAGGAAGAGGT (SEQ ID NO: 842)	TATGAACCACTAAGGCGTccaccttgttgGCCGGAAGCGAGGAAGAGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2415)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCTCTTCCTCGCTTCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3988)	1	EME2_-_1823220.23-P1P2	EME2	P1P2	na	na	na	na	na	na	na	na	na
GGCGCGGGTTGGACCCGGGA (SEQ ID NO: 843)	TATGAACCACTAAGGCGTccaccttgttgGGCGCGGGTTGGACCCGGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2416)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCCGGGTCCAACCCGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3989)	1	EME2_-_1823252.23-P1P2	EME2	P1P2	na	na	na	na	na	na	na	na	na
GCTAGGGACTGTGGCCTCGG (SEQ ID NO: 844)	TATGAACCACTAAGGCGTccaccttgttgGCTAGGGACTGTGGCCTCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2417)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGAGGCCACAGTCCCTAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3990)	1	ENDOV_-_78389318.23-P1P2	ENDOV	P1P2	na	na	na	na	na	na	na	na	na
GGAGAACCGCGCCCACCCCG (SEQ ID NO: 845)	TATGAACCACTAAGGCGTccaccttgttgGGAGAACCGCGCCCACCCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2418)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGGTGGGCGCGGTTCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3991)	1	ENDOV_+_78389220.23-P1P2	ENDOV	P1P2	na	na	na	na	na	na	na	na	na
GGCGGAAGCGGAGGGATCTC (SEQ ID NO: 846)	TATGAACCACTAAGGCGTccaccttgttgGGCGGAAGCGGAGGGATCTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2419)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGATCCCTCCGCTTCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3992)	1	ENDOV_-_78389210.23-P1P2	ENDOV	P1P2	na	na	na	na	na	na	na	na	na
GAGAGAGAGCGGGCGAGAGT (SEQ ID NO: 847)	TATGAACCACTAAGGCGTccaccttgttgGAGAGAGAGCGGGCGAGAGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2420)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACTCTCGCCCGCTCTCTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3993)	1	EP300_-_41487846.23-P1P2	EP300	P1P2	na	na	na	na	na	na	na	na	na
GCTACCGCTAGCCTGTAAGG (SEQ ID NO: 848)	TATGAACCACTAAGGCGTccaccttgttgGCTACCGCTAGCCTGTAAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2421)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTTACAGGCTAGCGGTAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3994)	1	EP300_-_41487756.23-P1P2	EP300	P1P2	na	na	na	na	na	na	na	na	na
GCTAGCCTGTAAGGAGGATT (SEQ ID NO: 849)	TATGAACCACTAAGGCGTccaccttgttgGCTAGCCTGTAAGGAGGATTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2422)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAATCCTCCTTACAGGCTAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3995)	1	EP300_-_41487762.23-P1P2	EP300	P1P2	na	na	na	na	na	na	na	na	na
GACGGCAGCGTCCAGGGTCC (SEQ ID NO: 850)	TATGAACCACTAAGGCGTccaccttgttgGACGGCAGCGTCCAGGGTCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2423)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGACCCTGGACGCTGCCGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 3996)	2	EP400_-_132434647.23-P1P2	EP400	P1P2	na	na	na	na	na	na	na	na	na

GCGCGTCGGCCGGTGA GGA (SEQ ID NO: 851)	TATGAAGCACTAAG3GCGTcaacttgtgG3GCTCG39CCGG	TCAGCAGCAGAGTTTAG3GAGTTAGGTCTTAAACTCCTGAC 3GCGG0393AGG03CAACAAGGH3GAG0CCTAGTGGTTCAT (SEQ ID NO: 3997)	2	EP400_+_13 2434562.23-P1P2	EP400	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGGCCGAGGGTAAG TGA (SEQ ID NO: 852)	TATGAACCACTAAGGCGTcaacttgtgG3GAGGCCAGGG	TCAACGAGCAGAGTTTAGGAGGTTAGGTCTTAAACTCACTTAG CCTGGGCCTCAGCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 3998)	2	EP400_-_ 132434651.23-P1P2	EP400	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGATGCAGTGTGGC GAA (SEQ ID NO: 853)	TATGAACCACTAAGGCGTcaacttgtgG3GCGATGCAGTG	TCAACGAGCAGAGTTTAGGAGGTTAGGTCTTAAACTTCCCCAG ACTGGATGGCGCCAAGAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 3999)	1	EP01_+_326 36107.23-P1P2	EP01	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGTCGCGGATGCAGT GTG (SEQ ID NO: 854)	TATGAACCACTAAGGCGTcaacttgtgCGGGTTGCGCCATG	TCAGCAGCAGAGTTTAGGAGGTTAGGTCTTAAACCCACACTGC ATCGACGCAGCGCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4000)	1	EP01_+_326 36112.23-P1P2	EP01	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAACAGCAGCGGGAG GGT (SEQ ID NO: 855)	TATGAACCACTAAGGCGTcaacttgtgTAACGACCAGCCG CGAG3G3TgttaagagctaagctgGGTAAAGTCTGCTGGTGTGA (SEQ ID NO: 4001)	TCAGCAGCAGAGTTTAGGAGGTTAGGTCTTAAACACCGTCCCC GGCTGGTGTCTTCAGCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4001)	1	EP01_-_ 32636004.2 3-P1P2	EP01	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAAGACTGCAGAGGGA TCG (SEQ ID NO: 856)	TATGAACCACTAAGGCGTcaacttgtgGAAAG3GACTGCAGA GGGATCGgttaagagctaagctgGGTAAAGTCTGCTGGTGTGA (SEQ ID NO: 4002)	TCAGCAGCAGAGTTTAGGAGGTTAGGTCTTAAACCGATCCCT CTGAAGTCTTTCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4002)	1	ERCC1_-_ 45927125.2 3-P1P2	ERCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
G3GACTGCAGAGAGGA ATCG (SEQ ID NO: 857)	TATGAACCACTAAG3GCGTcaacttgtgCGACTGCAGAG3G ATCGACGgttaagagctaagctgGGTAAAGTCTGCTGGTGTGA (SEQ ID NO: 4003)	TCAGCAGCAGAGTTTAG3GAGTTAGGTCTTAAACCCTCGATC CCTCTGCGAGTCCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4003)	1	ERCC1_-_ 45927128.2 3-P1P2	ERCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
G3TCCAGATGCTACGCCT CGG (SEQ ID NO: 858)	TATGAACCACTAAGGCGTcaacttgtgG3GTCCACGATGCTA IACGAATCTCGGACCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4004)	TCACCAGCAGAGTTTAGGAGGTTAGGTCTTAAACCCGAGGC 1 5928822.23-P1P2	1	ERCC1_+_4 5928822.23-P1P2	ERCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGCGCCATGAAGTGA GAA (SEQ ID NO: 859)	TATGAACCACTAAGGCGTcaacttgtgGCGCGCGCCATGAA TGATGGCGCGCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4005)	TCACGAGCAGAGTTTAGGAGGTTAGGTCTTAAACTTCTCACT 2 5873783.23-P1P2	2	ERCC2_+_4 5873783.23-P1P2	ERCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
G3GGAGCAGCAGTGAATA GCG (SEQ ID NO: 860)	TATGAAGCACTAAG3GCGTcaacttgtgCGGGAAG3AG3CAGT CTGCTGCTCCCCCAAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4006)	TCAGCAGCAGAGTTTAG3GAGTTAGGTCTTAAACCCGTATCA 2 5873641.23-P1P2	2	ERCC2_+_4 5873641.23-P1P2	ERCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGTCGTCGAGACCCAGA GAA (SEQ ID NO: 861)	TATGAACCACTAAG3GCGTcaacttgtgGGGTGTCGATCGCAGCC GCTCGACCACCCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4007)	TCACCAGCAGAGTTTAG3GAGCTTTAAACTTCTCTGTGG 2 4937357.6 2 3-P1P2	2	ERCC2_-_ 4937357.6 2 3-P1P2	ERCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCAAAGGGCTGTCCC AAC (SEQ ID NO: 862)	TATGAACCACTAAGGCGTcaacttgtgGCCCAAAG3GCCTG TCCCAAGgttaagagctaagctgGGTAAAGTCTGCTGGTGTGA (SEQ ID NO: 4008)	TCAGCAGCAGAGTTTAGGAGGTTAGGTCTTAAACGTTGGGA CAAGGCCTTTGG3GCCCAAGAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4008)	1	ERCC3_-_ 128051538 23-P1P2	ERCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
G3CGCTCCGCACAG3GCCCG CCG (SEQ ID NO: 863)	TATGAAGCACTAAG3GCGTcaacttgtgCGCGCTCCAGCAGG CCCGCCGCGgttaagagctaagctgGGTAAAGTCTGCTGGTGTGA (SEQ ID NO: 4009)	TCAGCAGCAGAGTTTAG3GAGTTAGGTCTTAAACCCGCGCG 1 CCTGTGGGAGCGGCCAAGAAGGTGGAGCGCCTAGTGGTTCAT A (SEQ ID NO: 4009)	1	ERCC3_-_ 128051729 23-P1P2	ERCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTCGACAACACGCGGAG3C TTC (SEQ ID NO: 864)	TATGAACCACTAAGGCGTcaacttgtgG3GTGGACAACACG3G CGAG3TTCgttaagagctaagctgGGTAAAGTCTGCTGGTGTGA (SEQ ID NO: 4010)	TCACCAGCAGAGTTTAGGAGGTTAGGTCTTAAACGGAAGGTC 1 CCTGTTGTCCACCCGGAAGAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4010)	1	ERCC3_+_1 28051740.23-P1P2	ERCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTTCAG3G3CAGCCGGCT CGA (SEQ ID NO: 865)	TATGAACCACTAAGGCGTcaacttgtgG3TGAGGGCGAGGC TGGCTGCCGCTGACCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4011)	TCACCAGCAGAGTTTAGGAGGTTAGGTCTTAAACTCGAG3CC 1 GGCTGCCGCTGACCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4011)	1	ERCC4_-_ 14014048.2 3-P1	ERCC4	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
G3GTCG3CTG3TACTCCAG CAG (SEQ ID NO: 866)	TATGAAGCACTAAG3GCGTcaacttgtgGGTCCGCTCG1AGT AGTACGAGGCGACCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4012)	TCAGCAGCAGAGTTTAG3GAGTTAGGTCTTAAACCTGGTGG 1 4014061.23-P1	1	ERCC4_+_1 4014061.23-P1	ERCC4	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
GTACTCGAGGAGGGGCG CCA (SEQ ID NO: 867)	TATGAACCACTAAGGCGTcaacttgtgGTACTCCAGCAGGC CCGCGCCAgttaagagctaagctgGGTAAAGTCTGCTGGTGTGA (SEQ ID NO: 4013)	TCACCAGCAGAGTTTAGGAGGTTAGGTCTTAAACTGGCGCC GCTCGCTGGAGGTACCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4013)	1	ERCC4_+_1 4014053.23-P1	ERCC4	P1	na	na	na	na	na	na	na	na	na	na	na	na	na

G9AGAGATCAGAGTAGT GGG (SEQ ID NO: 868)	TATGAAGCACTAAGAGGCGTcaacctgtgGAGACAGATCACAG TGTGATCTGTCCCAACAAG9TGGAGCGCCTTAGTGGTGATA (SEQ ID NO: 4014)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCCCACTAC	ERCC4_-_14025890.23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGAAGGCGGTGTGGTT GGT (SEQ ID NO: 869)	TATGAACCACTAAGAGCGTcaacctgtgGGGAGAGCGGTGT GTGTGGTGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4015)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCACAAGCA CACGGCCTCCCCCAACAAG9TGGAGCGCCTTAGTGGTTCATA	ERCC4_-_14025905.23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAGTGGGAGGGCGGTGT GGT (SEQ ID NO: 870)	TATGAACCACTAAGAGGCGTcaacctgtgGTAGTGGGAGCGCGG GTGTGGTGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4016)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCACAGCAGCG GCCTCCCCACTAGCAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4016)	ERCC4_-_14025902.23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTGGGGGGACGCGCACT AAT (SEQ ID NO: 871)	TATGAACCACTAAGAGGCGTcaacctgtgGAGTGGGGGACGCGG CACTAATgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4017)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCATTAGTGC CGTCCCCCACTCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4017)	ERCC5_+_1_03489446.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAGCGCAGGCGCTCA TGG (SEQ ID NO: 872)	TATGAACCACTAAGAGGCGTcaacctgtgGGGAGCGCAGGCTCA TCTCATGGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4018)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCCATGAG GCGGTCTGGTGTCCCCAACAAG9TGGAGCGCCTTAGTGGTTCAT A (SEQ ID NO: 4018)	ERCC5_-_103489622.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCCCCCACTGGAAAA CCG (SEQ ID NO: 873)	TATGAACCACTAAGAGGCGTcaacctgtgGGTCCCCCACTGG AAAACCGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4019)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCGGTTTTC CATGTGGGGGACCCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4019)	ERCC5_-_103489479.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGTCGAGCTGGGTT CCA (SEQ ID NO: 874)	TATGAAGCACTAAGAGCGTcaacctgtgCGCCGTCGAGCTG TGGTTCCAgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4020)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACTGCGACGC CAGCTCGACGGGCCCAACAAG9TGGAGCGCCTTAGTGGTTCAT A (SEQ ID NO: 4020)	ERCC6_+_5_07477060.23-P1P2	P1P2	ERCC6-P6B03_+_5_07477060.23-P1P2	P1P2	ERCC6-P6B03	P1P2	P6B03_+_5_07477060.23-P1P2	P1P2	na	na	na	na
GGTCGAAGCTGGGTTCCA ACG (SEQ ID NO: 875)	TATGAACCACTAAGAGGCGTcaacctgtgGGTCGAGCTGGGTT CCAACGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4021)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCGTTGGAA CCCCAGCTCGAACCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4021)	ERCC6_+_5_07477057.23-P1P2	P1P2	ERCC6-P6B03_+_5_07477057.23-P1P2	P1P2	ERCC6-P6B03	P1P2	P6B03_+_5_07477057.23-P1P2	P1P2	na	na	na	na
GTGGAGTTCCAAGGGCGC TGG (SEQ ID NO: 876)	TATGAACCACTAAGAGGCGTcaacctgtgGTGGAGTTCCAAGG GCGCCTGGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4022)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCCAGGCG CCCTTGGAAGCCCAACCAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4022)	ERCC6_+_5_07477050.23-P1P2	P1P2	ERCC6-P6B03_+_5_07477050.23-P1P2	P1P2	ERCC6-P6B03	P1P2	P6B03_+_5_07477050.23-P1P2	P1P2	na	na	na	na
G9AGCTTGGAGCTTAGA GTT (SEQ ID NO: 877)	TATGAAGCACTAAGAGCGTcaacctgtgGAGCTTGGAGTTAGA gttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4023)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACAACTCTAA GCTCCAAGCTCCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4023)	ERCC6L_+_71458843.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
G9AG9GCATCCCGAAGGT TTC (SEQ ID NO: 878)	TATGAACCACTAAGAGCGTcaacctgtgGAGAGCATCCCGA AGGTTTCgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4024)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAAGGAAACCTT CGGGAATGCCTCCCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4024)	ERCC6L_+_71458736.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTTAGCTAGACCCGGCC CAG (SEQ ID NO: 879)	TATGAAGCACTAAGAGGCGTcaacctgtgGGTTAGCTAGACC CGGCCCAGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4025)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCTGGGGCG GGTCTAGCTAACCCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4025)	ERCC6L_-_71458681.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCCCCTGAGCTCTG TCA (SEQ ID NO: 880)	TATGAAGCACTAAGAGCGTcaacctgtgCGCCGCCCCTGAGC TCTCTGTCAgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4026)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACTGACAGA GCTCAGGGGGGGCGGCCAACAAG9TGGAGCGCCTTAGTGGTTCAT A (SEQ ID NO: 4026)	ERCC6_+_60240671.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCATATCGTGTCGTCAC AC (SEQ ID NO: 881)	TATGAACCACTAAGAGGCGTcaacctgtgGGCATATCGTGTCG TCACACgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4027)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACGGTGTGTAG GACAGGATATGGCCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4027)	ERCC6_-_80240352.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTGAGGACACGATATG CTG (SEQ ID NO: 882)	TATGAAGCACTAAGAGGCGTcaacctgtgGGTGAGGACACGA CTGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4028)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACCAGCATAT CGTGTCGTCACCCAAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4028)	ERCC6_+_60240826.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGGTCAGACTACGGC CAG (SEQ ID NO: 883)	TATGAAGCACTAAGAGCGTcaacctgtgGAGGGGTCAGACTA CGGCCCAGgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4029)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACTCTGGGTGC TAGTCTGACCCTCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4029)	ESCO1_-_19160322.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCGGCGACGGCAAAA G9A (SEQ ID NO: 884)	TATGAACCACTAAGAGGCGTcaacctgtgGCCTCGGCGACGGC AAAAGG9AgttaagagctaagtgcGGTAAAGTCTGGTGGTGA (SEQ ID NO: 4030)	TCAGCAGCAGAGCTTTAGCGCAGCTTAGCTCTTAAACGTCGTTTTG CCGTCGCG9GAGCGCCAACAAG9TGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4030)	ESCO1_+_1_91980582.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGGTACGCGAGCGTCCCTGA (SEQ ID NO: 885)	TATGAACCACTAAGGCGTccaccttgttgGGGTACGCGAGCGTCCCTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2458)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCAGGGACGCTCGCGTACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4031)	1		ESCO1_+_19180562.23-P1P2	ESCO1	P1P2	na	na	na	na	na	na	na	na	na
GGACTGCCCCGGGCCGCTGA (SEQ ID NO: 886)	TATGAACCACTAAGGCGTccaccttgttgGGACTGCCCCGGGCCGCTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2459)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCAGCGGCCCGGGGCAGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4032)	2		ESPL1_-_53662296.23-P1P2	ESPL1	P1P2	na	na	na	na	na	na	na	na	na
GGTCCTGTACCTAGGAAAGA (SEQ ID NO: 887)	TATGAACCACTAAGGCGTccaccttgttgGGTCCTGTACCTAGGAAAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2460)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTTTCCTAGGTACAGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4033)	2		ESPL1_-_53662146.23-P1P2	ESPL1	P1P2	na	na	na	na	na	na	na	na	na
GTCCTGGAGGGAGTGTCGGG (SEQ ID NO: 888)	TATGAACCACTAAGGCGTccaccttgttgGTCCTGGAGGGAGTGTCGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2461)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGACACTCCCTCCAGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4034)	2		ESPL1_-_53662268.23-P1P2	ESPL1	P1P2	na	na	na	na	na	na	na	na	na
GCGTCCACGGGTGAGTATGG (SEQ ID NO: 889)	TATGAACCACTAAGGCGTccaccttgttgGCGTCCACGGGTGAGTATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2462)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATACTCACCCGTGGACGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4035)	1		EWSR1_-_29664350.23-P1P2	EWSR1	P1P2	na	na	na	na	na	na	na	na	na
GGTTCCACCATACTCACCCG (SEQ ID NO: 890)	TATGAACCACTAAGGCGTccaccttgttgGGTTCCACCATACTCACCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2463)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGTGAGTATGGTGGAACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4036)	1		EWSR1_+_29664332.23-P1P2	EWSR1	P1P2	na	na	na	na	na	na	na	na	na
GTGGCGTCCACGGGTGAGTA (SEQ ID NO: 891)	TATGAACCACTAAGGCGTccaccttgttgGTGGCGTCCACGGGTGAGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2464)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACTCACCCGTGGACGCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4037)	1		EWSR1_-_29664347.23-P1P2	EWSR1	P1P2	na	na	na	na	na	na	na	na	na
GGGAACCCGCGCAAATTGAA (SEQ ID NO: 892)	TATGAACCACTAAGGCGTccaccttgttgGGGAACCCGCGCAAATTGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2465)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCAATTTGCGCGGGTTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4038)	1		EXO1_-_242011502.23-P1P2	EXO1	P1P2	na	na	na	na	na	na	na	na	na
GGGCCGGCAACGCAGAACAC (SEQ ID NO: 893)	TATGAACCACTAAGGCGTccaccttgttgGGGCCGGCAACGCAGAACACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2466)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTGTTCTGCGTTGCCGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4039)	1		EXO1_+_242011534.23-P1P2	EXO1	P1P2	na	na	na	na	na	na	na	na	na
GTGGCCAGAGCACCCACGGC (SEQ ID NO: 894)	TATGAACCACTAAGGCGTccaccttgttgGTGGCCAGAGCACCCACGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2467)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCGTGGGTGCTCTGGCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4040)	1		EXO1_+_242011550.23-P1P2	EXO1	P1P2	na	na	na	na	na	na	na	na	na
GACCCGCGGCTCGGGATGGA (SEQ ID NO: 895)	TATGAACCACTAAGGCGTccaccttgttgGACCCGCGGCTCGGGATGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2468)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCATCCCGAGCCGCGGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4041)	1		EZH1_+_40897027.23-P1P2	EZH1	P1P2	na	na	na	na	na	na	na	na	na
GCGGGACCCGCGGCTCGGGA (SEQ ID NO: 896)	TATGAACCACTAAGGCGTccaccttgttgGCGGGACCCGCGGCTCGGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2469)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCCGAGCCGCGGGTCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4042)	1		EZH1_+_40897031.23-P1P2	EZH1	P1P2	na	na	na	na	na	na	na	na	na
GGTCCTAGCAGCGGGACCCG (SEQ ID NO: 897)	TATGAACCACTAAGGCGTccaccttgttgGGTCCTAGCAGCGGGACCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2470)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGTCCCGCTGCTAGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4043)	1		EZH1_+_40897041.23-P1P2	EZH1	P1P2	na	na	na	na	na	na	na	na	na
GCCGGACCGAGCGCCAAACG (SEQ ID NO: 898)	TATGAACCACTAAGGCGTccaccttgttgGCCGGACCGAGCGCCAAACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2471)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTTTGGCGCTCGGTCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4044)	1		EZH2_-_148581386.23-P1P2	EZH2	P1P2	na	na	na	na	na	na	na	na	na
GGCCCCGCTCGGCGATACCC (SEQ ID NO: 899)	TATGAACCACTAAGGCGTccaccttgttgGGCCCCGCTCGGCGATACCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2472)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGTATCGCCGAGCGGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4045)	1		EZH2_-_148581155.23-P1P2	EZH2	P1P2	na	na	na	na	na	na	na	na	na
GGTCGCGTCCGACACCCGGT (SEQ ID NO: 900)	TATGAACCACTAAGGCGTccaccttgttgGGTCGCGTCCGACACCCGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2473)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCGGGTGTCGGACGCGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4046)	1		EZH2_+_148581344.23-P1P2	EZH2	P1P2	na	na	na	na	na	na	na	na	na
GATGGAGGGGGAGAGTACGT (SEQ ID NO: 901)	TATGAACCACTAAGGCGTccaccttgttgGATGGAGGGGGAGAGTACGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2474)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGTACTCTCCCCCTCCATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4047)	1		FAM175A_+_84406203.23-P1P2	FAM175A	P1P2	na	na	na	na	na	na	na	na	na

GCCTCTTGTGTAGCCTGAGG (SEQ ID NO: 902)	TATGAACCACTAAGGCGTccaccttgttgGCCTCTTGTGTAGCCTGAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2475)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTCAGGCTACACAAGAGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4048)	1		FAM175A_+_84406232.23-P1P2	FAM175A	P1P2	na	na	na	na	na	na	na	na	na
GGAGGCGGCGGTAGCATGGA (SEQ ID NO: 903)	TATGAACCACTAAGGCGTccaccttgttgGGAGGCGGCGGTAGCATGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2476)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCATGCTACCGCCGCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4049)	1		FAM175A_+_84406217.23-P1P2	FAM175A	P1P2	na	na	na	na	na	na	na	na	na
GACAGGGCACCGAGGGAAGG (SEQ ID NO: 904)	TATGAACCACTAAGGCGTccaccttgttgGACAGGGCACCGAGGGAAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2477)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTTCCCTCGGTGCCCTGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4050)	1		FAN1_-_31196149.23-P1P2	FAN1	P1P2	na	na	na	na	na	na	na	na	na
GCGAGGGAAGGAGGACGCGA (SEQ ID NO: 905)	TATGAACCACTAAGGCGTccaccttgttgGCGAGGGAAGGAGGACGCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2478)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCGTCCTCCTTCCCTCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4051)	1		FAN1_-_31196158.23-P1P2	FAN1	P1P2	na	na	na	na	na	na	na	na	na
GGTGACAGGGCACCGAGGGA (SEQ ID NO: 906)	TATGAACCACTAAGGCGTccaccttgttgGGTGACAGGGCACCGAGGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2479)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCCTCGGTGCCCTGTCACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4052)	1		FAN1_-_31196146.23-P1P2	FAN1	P1P2	na	na	na	na	na	na	na	na	na
GAAGGCCATGTCCGACTCGT (SEQ ID NO: 907)	TATGAACCACTAAGGCGTccaccttgttgGAAGGCCATGTCCGACTCGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2480)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGAGTCGGACATGGCCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4053)	1		FANCA_+_89883007.23-P1P2	FANCA	P1P2	na	na	na	na	na	na	na	na	na
GGAGTTCGGGACCCACGAGT (SEQ ID NO: 908)	TATGAACCACTAAGGCGTccaccttgttgGGAGTTCGGGACCCACGAGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2481)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACTCGTGGGTCCCGAACTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4054)	1		FANCA_-_89883018.23-P1P2	FANCA	P1P2	na	na	na	na	na	na	na	na	na
GGGGACCCACGAGTCGGACA (SEQ ID NO: 909)	TATGAACCACTAAGGCGTccaccttgttgGGGGACCCACGAGTCGGACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2482)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTCCGACTCGTGGGTCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4055)	1		FANCA_-_89883024.23-P1P2	FANCA	P1P2	na	na	na	na	na	na	na	na	na
GCAACGCAGCGGCAACATAC (SEQ ID NO: 910)	TATGAACCACTAAGGCGTccaccttgttgGCAACGCAGCGGCAACATACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2483)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTATGTTGCCGCTGCGTTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4056)	1		FANCB_-_14891130.23-P1P2	FANCB	P1P2	na	na	na	na	na	na	na	na	na
GCGGAGGCCCCACGTCGATA (SEQ ID NO: 911)	TATGAACCACTAAGGCGTccaccttgttgGCGGAGGCCCCACGTCGATAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2484)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTATCGACGTGGGGCCTCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4057)	1		FANCB_-_14891149.23-P1P2	FANCB	P1P2	na	na	na	na	na	na	na	na	na
GGCGTCGAACGCAGAGGAGG (SEQ ID NO: 912)	TATGAACCACTAAGGCGTccaccttgttgGGCGTCGAACGCAGAGGAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2485)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTCCTCTGCGTTCGACGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4058)	1		FANCB_-_14891231.23-P1P2	FANCB	P1P2	na	na	na	na	na	na	na	na	na
GGAGGGCGACCGGCTCAAAG (SEQ ID NO: 913)	TATGAACCACTAAGGCGTccaccttgttgGGAGGGCGACCGGCTCAAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2486)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTTGAGCCGGTCGCCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4059)	1		FANCC_-_98079365.23-P1P2	FANCC	P1P2	na	na	na	na	na	na	na	na	na
GGGTCTGGCCGACACGTCAG (SEQ ID NO: 914)	TATGAACCACTAAGGCGTccaccttgttgGGGTCTGGCCGACACGTCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2487)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGACGTGTCGGCCAGACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4060)	1		FANCC_+_98079737.23-P1P2	FANCC	P1P2	na	na	na	na	na	na	na	na	na
GTCCCGCGGTCGCCCGGCAG (SEQ ID NO: 915)	TATGAACCACTAAGGCGTccaccttgttgGTCCCGCGGTCGCCCGGCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2488)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGCCGGGCGACCGCGGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4061)	1		FANCC_-_98079954.23-P1P2	FANCC	P1P2	na	na	na	na	na	na	na	na	na
GCGGGAAAGTCGAAAACTAC (SEQ ID NO: 916)	TATGAACCACTAAGGCGTccaccttgttgGCGGGAAAGTCGAAAACTACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2489)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTAGTTTTCGACTTTCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4062)	1		FANCD2_-_10068140.23-P1P2	FANCD2	P1P2	na	na	na	na	na	na	na	na	na
GCGGTGAGTAAGTGGAGCAA (SEQ ID NO: 917)	TATGAACCACTAAGGCGTccaccttgttgGCGGTGAGTAAGTGGAGCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2490)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGCTCCACTTACTCACCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4063)	1		FANCD2_-_10068176.23-P1P2	FANCD2	P1P2	na	na	na	na	na	na	na	na	na
GGAAAGTCGAAAACTACGGG (SEQ ID NO: 918)	TATGAACCACTAAGGCGTccaccttgttgGGAAAGTCGAAAACTACGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2491)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCGTAGTTTTCGACTTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4064)	1		FANCD2_-_10068143.23-P1P2	FANCD2	P1P2	na	na	na	na	na	na	na	na	na

GGAACGGCTGCGGCTTCGGG (SEQ ID NO: 919)	TATGAACCACTAAGGCGTccaccttgttgGGAACGGCTGCGGCTTCGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2492)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGAAGCCGCAGCCGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4065)	1	FANCE_-_35420157.23-P1P2	FANCE	P1P2	na	na	na	na	na	na	na	na	na
GGACAGCGCGGGAACGGCTG (SEQ ID NO: 920)	TATGAACCACTAAGGCGTccaccttgttgGGACAGCGCGGGAACGGCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2493)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGCCGTTCCCGCGCTGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4066)	1	FANCE_-_35420147.23-P1P2	FANCE	P1P2	na	na	na	na	na	na	na	na	na
GTTGGGAGACCGGAGAAACC (SEQ ID NO: 921)	TATGAACCACTAAGGCGTccaccttgttgGTTGGGAGACCGGAGAAACCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2494)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGTTTCTCCGGTCTCCCAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4067)	1	FANCE_+_35420158.23-P1P2	FANCE	P1P2	na	na	na	na	na	na	na	na	na
GCTGCAGAAGGGATTCCATG (SEQ ID NO: 922)	TATGAACCACTAAGGCGTccaccttgttgGCTGCAGAAGGGATTCCATGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2495)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCATGGAATCCCTTCTGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4068)	1	FANCF_-_22647359.23-P1P2	FANCF	P1P2	na	na	na	na	na	na	na	na	na
GGCGGCTGCACAACCAGTGG (SEQ ID NO: 923)	TATGAACCACTAAGGCGTccaccttgttgGGCGGCTGCACAACCAGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2496)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACTGGTTGTGCAGCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4069)	1	FANCF_+_22647155.23-P1P2	FANCF	P1P2	na	na	na	na	na	na	na	na	na
GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	TATGAACCACTAAGGCGTccaccttgttgGGGGTCCCAGGTGCTGACGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2497)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGTCAGCACCTGGGACCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4070)	1	FANCF_-_22647294.23-P1P2	FANCF	P1P2	na	na	na	na	na	na	na	na	na
GCGCAGACCCGGCCAACCAG (SEQ ID NO: 925)	TATGAACCACTAAGGCGTccaccttgttgGCGCAGACCCGGCCAACCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2498)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGGTTGGCCGGGTCTGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4071)	1	FANCG_+_35079731.23-P1P2	FANCG	P1P2	na	na	na	na	na	na	na	na	na
GGCCCGCTCGGCTCTCGCGG (SEQ ID NO: 926)	TATGAACCACTAAGGCGTccaccttgttgGGCCCGCTCGGCTCTCGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2499)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCGAGAGCCGAGCGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4072)	1	FANCG_-_35079920.23-P1P2	FANCG	P1P2	na	na	na	na	na	na	na	na	na
GTCTCAACCTCGCCTCTGGT (SEQ ID NO: 927)	TATGAACCACTAAGGCGTccaccttgttgGTCTCAACCTCGCCTCTGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2500)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCAGAGGCGAGGTTGAGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4073)	1	FANCG_-_35079741.23-P1P2	FANCG	P1P2	na	na	na	na	na	na	na	na	na
GAGCCGCTCTGCAGTCACCT (SEQ ID NO: 928)	TATGAACCACTAAGGCGTccaccttgttgGAGCCGCTCTGCAGTCACCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2501)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGGTGACTGCAGAGCGGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4074)	1	FANCI_+_89787270.23-P1P2	FANCI	P1P2	na	na	na	na	na	na	na	na	na
GGAACCGCGAAAGCTGAGGC (SEQ ID NO: 929)	TATGAACCACTAAGGCGTccaccttgttgGGAACCGCGAAAGCTGAGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2502)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCTCAGCTTTCGCGGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4075)	1	FANCI_+_89787428.23-P1P2	FANCI	P1P2	na	na	na	na	na	na	na	na	na
GGTGGCGGCGTTGGGTTGAG (SEQ ID NO: 930)	TATGAACCACTAAGGCGTccaccttgttgGGTGGCGGCGTTGGGTTGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2503)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCAACCCAACGCCGCCACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4076)	1	FANCI_-_89787239.23-P1P2	FANCI	P1P2	na	na	na	na	na	na	na	na	na
GACCGCCATGGCTCGAAGTC (SEQ ID NO: 931)	TATGAACCACTAAGGCGTccaccttgttgGACCGCCATGGCTCGAAGTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2504)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGACTTCGAGCCATGGCGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4077)	1	FANCL_-_58468461.23-P1P2	FANCL	P1P2	na	na	na	na	na	na	na	na	na
GCTTCGAGCCATGGCGGTGA (SEQ ID NO: 932)	TATGAACCACTAAGGCGTccaccttgttgGCTTCGAGCCATGGCGGTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2505)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCACCGCCATGGCTCGAAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4078)	1	FANCL_+_58468435.23-P1P2	FANCL	P1P2	na	na	na	na	na	na	na	na	na
GTCCGGACTTCGAGCCATGG (SEQ ID NO: 933)	TATGAACCACTAAGGCGTccaccttgttgGTCCGGACTTCGAGCCATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2506)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATGGCTCGAAGTCCGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4079)	1	FANCL_+_58468441.23-P1P2	FANCL	P1P2	na	na	na	na	na	na	na	na	na
GATGGGGATCGGAACCGTAG (SEQ ID NO: 934)	TATGAACCACTAAGGCGTccaccttgttgGATGGGGATCGGAACCGTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2507)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTACGGTTCCGATCCCCATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4080)	1	FANCM_-_45605178.23-P1P2	FANCM	P1P2	na	na	na	na	na	na	na	na	na
GGAAGGAAACCGATGGGGAT (SEQ ID NO: 935)	TATGAACCACTAAGGCGTccaccttgttgGGAAGGAAACCGATGGGGATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2508)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATCCCCATCGGTTTCCTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4081)	1	FANCM_-_45605167.23-P1P2	FANCM	P1P2	na	na	na	na	na	na	na	na	na

GTGGTTGTGGGCCTAAT GAG (SEQ ID NO: 936)	TATGAAGCACTAAGGCGTcaacctgtgGTGGTTGTGGGCC TAATGAAgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2509)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCTCATTAG GCCGACAACCACCAACAAGGTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4082)	FANCM_-45605241.2 3-P1P2	FANCM	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTCTGGGGAGGAAGT CGG (SEQ ID NO: 937)	TATGAACCACTAAGGCGTcaacctgtgGAGCTCGGGAGG AAAGTCGgttaagagtaaagtcgtGTAAAGTCTGGTGGTGA (SEQ ID NO: 2510)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCGAGTCG CTCGGCGAGCGTCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4083)	FBXO18_-59321882.23-P1P2	FBXO18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGCAGCCGACGTCCTC CTC (SEQ ID NO: 938)	TATGAACCACTAAGGCGTcaacctgtgGCCGGCAGCCGA CGCTCCTCgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2511)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCGAGAAGG AGCTCGCTCGCGCGAAGAAGTCGCGACGCCTTAGTCGTTCATA (SEQ ID NO: 4084)	FBXO18_+_5932212.23-P1P2	FBXO18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGCTGCCGGGGGAC GCT (SEQ ID NO: 939)	TATGAACCACTAAGGCGTcaacctgtgGCTCGCTGCCGG GGACGCTgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2512)	CACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCACGGTCG CCGGGCAGCGACGAAGGTGGACGCCTTAGTCGTTCATA (SEQ ID NO: 4085)	FBXO18_-59302945.23-P1P2	FBXO18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGACGCCCGGACAGAG CGA (SEQ ID NO: 940)	TATGAACCACTAAGGCGTcaacctgtgGAGACGCCCGGA CAGAGCGAgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2513)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCTCGCTCTG TGCCGGCGCTCTCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4086)	FBXW7_+_1534572062.2-ENST00000269718.4	FBXW7	ENST0000026917184	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCCTCACTCCAGAGA GAG (SEQ ID NO: 941)	TATGAACCACTAAGGCGTcaacctgtgGCGCCTCACTCCA GAGGAGAGgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2514)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCTCTCTCT GGAGTGAGGCGCGCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4087)	FBXW7_-_1534572482-3-ENST00000026923.4	FBXW7	ENST0000026923	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTGAGACGCCAAGGTCT CGG (SEQ ID NO: 942)	TATGAACCACTAAGGCGTcaacctgtgGGTGAGACGCAAG cTTCGCGTCTCACCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4088)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCTCGACAC GCTTCGCGTCTCACCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4088)	FBXW7_-_1534571956.2-ENST00000026923.4	FBXW7	ENST0000026923	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGCAGCTGCGGAAGG CTC (SEQ ID NO: 943)	TATGAACCACTAAGGCGTcaacctgtgGCGGCAGCTGCGG AAAGGCTCgttaagagtaaagtcgtGTAAAGTCTGGTGGTGA (SEQ ID NO: 2515)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACGAAGCCTT CCGCGCAGCTGCGCGCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4089)	FBXW7_-_1534581777.2-ENST00000269718.4	FBXW7	ENST0000026917184	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGCGGGCGCGAGAAT AGT (SEQ ID NO: 944)	TATGAACCACTAAGGCGTcaacctgtgGGGCGGCGGGCG CGCGAGAAgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2517)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACGCTTTCTC GGGCGCCGCGGGCCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4090)	FBXW7_-_1534561277.2-ENST00000269718.4	FBXW7	ENST0000026917184	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCCAGGGGCCACTTCT CCC (SEQ ID NO: 945)	TATGAAGCACTAAGGCGTcaacctgtgGGTCCAGGGGCC ACTTCTCCCgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2518)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACGTGGACAA GGTCCAGGGGCCGCCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4091)	FBXW7_-_1534553467.2-ENST00000269718.4	FBXW7	ENST0000026917184	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATCTACCTGGCTTGAG GAA (SEQ ID NO: 946)	TATGAACCACTAAGGCGTcaacctgtgGATCTACCTGGCTT GAGGAAgttaagagtaaagtcgtGTAAAGTCTGGTGGTGA (SEQ ID NO: 2519)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACTTCCTCAA GCCACGGTACATCCAAACAAGGTCGACGCCCTTAGTCGTTCATA (SEQ ID NO: 4092)	FBXW7_-_1533328899.2-ENST00000269718.4	FBXW7	ENST0000026917184	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAAAAACGACGACGAAC TGG (SEQ ID NO: 947)	TATGAACCACTAAGGCGTcaacctgtgGGCAAAAACGACG ACGAACTGgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2520)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCCCAGTTC GTCGTCTTTTGCCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4093)	FBXW7_+_1533328955.23-ENST00000269718.4	FBXW7	ENST0000026917184	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAGGAAAAGACGACG AAC (SEQ ID NO: 948)	TATGAACCACTAAGGCGTcaacctgtgCGGCAGGAAAAG ACGAACGAgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2521)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACGTTCCGTC GTCTTTTGCCTGCCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4094)	FBXW7_+_1533328955.23-ENST00000269718.4	FBXW7	ENST0000026917184	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGGAGCGCGGGCTTT CGT (SEQ ID NO: 949)	TATGAACCACTAAGGCGTcaacctgtgGCGCGGAGCGCG GGCTTTCGgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2522)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACGTCGAAAGC GCGCGCTCGCCCGCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4095)	FEN1_-_81560382.23-P1P2	FEN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGAAGGGAAGAGCGAG CTT (SEQ ID NO: 950)	TATGAACCACTAAGGCGTcaacctgtgGGAGAAGGGAAG AGCGAGCTgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2523)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACAAGCCTCG CTCTCGCCTTCTCGCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4096)	FEN1_-_81560434.23-P1P2	FEN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGCGTGCGCGCTTTCGA AGG (SEQ ID NO: 951)	TATGAAGCACTAAGGCGTcaacctgtgGTGAGCGTGCGC GCTTTCGAgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2524)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCCTTCCAA AGCCCGGCGCTCCCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4097)	FEN1_-_81560383.23-P1P2	FEN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CAGGTAAGAGTGTGCTCTA TGG (SEQ ID NO: 952)	TATGAACCACTAAGGCGTcaacctgtgCAGGTAAGAGTG TCTATGCTgttaagagtaaagtcgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2525)	TCACCAGCAGAGACTTTACGGCAGCTTAGCTCTTAAACCCATAGA GGCCACTCTTCCTCCCAACAAGGTCGACGCTTAGTCGTTCATA (SEQ ID NO: 4098)	PTPL1_-_54244091.23-P1P2	PTPL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCGAGACTCCAATCCCA GCA (SEQ ID NO: 953)	TATGAACCACTAAGGCGTcaactgtgGCGAGACTCCAATC (SEQ ID NO: 4099)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACTCCTGGG ATTCGAGTCTCCGCCAACAAGTGTGGACCGCCTAGTGGTCATA	2	FIP1L1_+_64243863.23-P1P2	FIP1L1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCGTGTCCGCGCTGA CTT (SEQ ID NO: 954)	TATGAACCACTAAGGCGTcaactgtgGCGTCGTGTCCGCGC (SEQ ID NO: 2527)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACAAGTCAG GCCGTGACAGCGACCCAACAAGGTGGACGCCTTAGTGGTTCAT A	2	FIP1L1_-_54244223.23-P1P2	FIP1L1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACCAAAAGACTACGTCG	TATGAACCACTAAGGCGTcaactgtgAAGCAAAAGACTA (SEQ ID NO: 2528)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGGGGAAGC TAGTCTTTGGTTCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GEN1_+_179635450.23-P1P2	GEN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATCAGTCGCGACGACCC GGG (SEQ ID NO: 955)	TATGAACCACTAAGGCGTcaactgtgGATCAGTGGGACG (SEQ ID NO: 2529)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACCCCGGGT CCTCCCACTGATCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GEN1_+_179635942.23-P1P2	GEN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCTCAACCGCGCAGT CCC (SEQ ID NO: 957)	TATGAACCACTAAGGCGTcaactgtgGGGGCTCAACCGC (SEQ ID NO: 2530)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGGGCCTG GCGGTTGAGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GEN1_+_179635431.23-P1P2	GEN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGAGGATTTGGAGGG TAC (SEQ ID NO: 958)	TATGAACCACTAAGGCGTcaactgtgGAAGAGGATTTGG (SEQ ID NO: 2531)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGTACCCTC CAAATCCTCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GPN1_-_27952137.23-P1P2	GPN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCCCCGAAAGCCATG CGC (SEQ ID NO: 959)	TATGAACCACTAAGGCGTcaactgtgCGCCCCCGAAAGC (SEQ ID NO: 2532)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGTCGCATG GCTTTCGGCGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GPN1_+_27852134.23-P1P2	GPN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACGAAGAATGCAGAGGC GGG (SEQ ID NO: 961)	TATGAACCACTAAGGCGTcaactgtgGACCGAAGAAATgtca (SEQ ID NO: 2534)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACCCCGCCT GCATTCTTCGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H1_-_183443918.23-P1P2	GTF2H1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCCTAGTAACAGAGGT CGG (SEQ ID NO: 962)	TATGAACCACTAAGGCGTcaactgtgCCCCCTAGTAACAGG (SEQ ID NO: 2535)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACCCGGGTCCT GTTACTAGGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H1_-_183441682.23-P1P2	GTF2H1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGAGAGCCCCTAGTAA CAG (SEQ ID NO: 963)	TATGAACCACTAAGGCGTcaactgtgCGCGAGAGCCCCTA (SEQ ID NO: 2536)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACCCTGTTAGT GTAACACGcgttaaagcgtaaagcgtaaagcgtaaagcgtaaagcgt AGGGGTGTCTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H1_-_183441601.23-P1P2	GTF2H1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGAAGGTGGCGCAGGCC GTC (SEQ ID NO: 964)	TATGAACCACTAAGGCGTcaactgtgGAGAGGTGGCGCA (SEQ ID NO: 2537)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGACGCGCC TGCCCCACCTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H2_+_70383428.23-ENST00000027440.5:ENST00000033028.7:4400.5:ENS	GTF2H2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGTCTGCCTACGCT AGG (SEQ ID NO: 965)	TATGAACCACTAAGGCGTcaactgtgCGGCGCTCTGTGGCT (SEQ ID NO: 2538)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACCCTACGCTA GCCAGACGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H2_+_70383416.23-ENST00000027440.5:ENST00000033028.7:4400.5:ENS	GTF2H2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCTGTGAGGCGCGAGAG AGAGCT (SEQ ID NO: 966)	TATGAACCACTAAGGCGTcaactgtgGTCTGTGAGGCGC (SEQ ID NO: 2539)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGCACGTCTG CGCCTCACAGACGCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H2_+_70383440.23-ENST00000027440.5:ENST00000033028.7:4400.5:ENS	GTF2H2	P1P2	GTF2H2D	GTF2H2D_-_68956109.2-3-P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGACGGTGAGGCACCT GCA (SEQ ID NO: 967)	TATGAACCACTAAGGCGTcaactgtgAGACGGTGAGGCA (SEQ ID NO: 2540)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACTGCAGGG TCCTCACCGTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H3_-_124119435.23-P1P2	GTF2H3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGACGAACACAGCCGAG CCC (SEQ ID NO: 968)	TATGAACCACTAAGGCGTcaactgtgCGAACCGACGAGCA (SEQ ID NO: 2541)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGTCGCCTCC GCGTGTCTCGGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H3_+_124118448.23-P1P2	GTF2H3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTTCGAAGTCCCGCGCCT GCA (SEQ ID NO: 969)	TATGAACCACTAAGGCGTcaactgtgGGTTCGAAGTCCGCG (SEQ ID NO: 2542)	TCACCAGCAGAGCTTACGCAGCTTAGCTCTTAAACTGCAGGG CGGGGACTTCGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	GTF2H3_+_124118426.23-P1P2	GTF2H3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GACACTCGGGGTGTGCGG AG3 (SEQ ID NO: 970)	TATGAACCACTAAGGCGTcaacctgtgGAGACTCGGGGTG CACCCGAGTGTCCACAAGGGTCGACGCTTAGTGGTGATCATA (SEQ ID NO: 2543)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCCTCCGG 308781672 3-P1P2	GTF2H4_308781672	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGATCGAACAGATG GTG (SEQ ID NO: 971)	TATGAACCACTAAGGCGTcaacctgtgGAGGATCGAACAG AGTCGTGTGgttaagagctaagctgtGTAAAGTCTGGTGGTGA (SEQ ID NO: 2544)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCACCACTC TGTGGATCCTCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4117)	GTF2H4_308781652 3-P1P2	GTF2H4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CGCCGACCAAGAATCGACA GAG (SEQ ID NO: 972)	TATGAACCACTAAGGCGTcaacctgtgGAGCAAGAAGTCTGCTGGTGA (SEQ ID NO: 2545)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCTCCTGTGTCG ATCCTCCTCGTCGCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4118)	GTF2H4_308781602 3-P1P2	GTF2H4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATTCTGTCGGCCCGAGGT GTG (SEQ ID NO: 973)	TATGAACCACTAAGGCGTcaacctgtgGATCTGTCGGCCG AGCCCACGAGATCCAACAAGGTTGGACGCTTAGTGGTTCATA (SEQ ID NO: 2546)	CACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCACACCTC GGCCCACGAGATCCAACAAGGTTGGACGCTTAGTGGTTCATA (SEQ ID NO: 4119)	GTF2H5_158689434 23-P1P2	GTF2H5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCACTGCTCGCACACA CCT (SEQ ID NO: 974)	TATGAACCACTAAGGCGTcaacctgtgGGGCACGCTCGCCA CACACCTgttaagagctaagctgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2547)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCACGCTGTG TGGCGAGCGCTGTCCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4120)	GTF2H5_895689422 23-P1P2	GTF2H5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGTGTGGCGGAGCGTCC GAG (SEQ ID NO: 975)	TATGAACCACTAAGGCGTcaacctgtgGTGTGTGGCGGAGC GTCCGAGgttaagagctaagctgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2548)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCTCGGGAC GCTCGGCCACACAGCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4121)	GTF2H5_158689449 23-P1P2	GTF2H5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGCGAGTCACACATG AG3 (SEQ ID NO: 976)	TATGAAGCACTAAGGCGTcaacctgtgCGCGCGCGAGTCAC CATGAGgttaagagctaagctgtGCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2549)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCCTCATG GTGACTCGGGCGGCCAACAAGGTTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4122)	GTF3C4_136546125 23-P1P2	GTF3C4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGAGAAAAACCGCGCGC CGA (SEQ ID NO: 977)	TATGAACCACTAAGGCGTcaacctgtgGGGGAGAAAAACCG CCGCGTGAgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2550)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCTCCGCGCG CGGTTTTCTCCCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4123)	GTF3C4_-1365456899 23-P1P2	GTF3C4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGGGGGAGAAAAACCG CCG (SEQ ID NO: 978)	TATGAAGCACTAAGGCGTcaacctgtgGTGAGGGGAGAAA ACCGCCGTgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2551)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACCCGCGGT TTTCGTCCCCTCACCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4124)	GTF3C4_-1365456885 23-P1P2	GTF3C4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACATGCTAGCGAGAGTA GAC (SEQ ID NO: 979)	TATGAAGCACTAAGGCGTcaacctgtgCGATATGCTAGCGA GTGTAGACgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2552)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACGTCTAGCT CGCTAGCATGTCCAAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4125)	HZAFX_-1186866121 23-P1P2	HZAFX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGGCGCGCGAAGAGC TGG (SEQ ID NO: 980)	TATGAACCACTAAGGCGTcaacctgtgGCGGGGCGCGCGA AGACTCGGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2553)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACGCCAGTCTT GCCGCCGGCCCGCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4126)	HZAFX_+11-1396680723 23-P1P2	HZAFX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTAGAGCGGTGAAGCA CGA (SEQ ID NO: 981)	TATGAACCACTAAGGCGTcaacctgtgGGTAGAGCGGTGA AGCACGAgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2554)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACTCCGTGCTT CACGCGTCTCACCCAAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4127)	HZAFX_-1186866134 23-P1P2	HZAFX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAATCCGGACTTCCCGGATT GA (SEQ ID NO: 982)	TATGAAGCACTAAGGCGTcaacctgtgGAATCCGGACTTCC GGATTGAgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2555)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACGTCATCCG GGATAACTCTGGATTTCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4128)	HZAFZ_+10-1087147523 P1P2	HZAFZ	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGTGGTGGGCAGCCATT GG (SEQ ID NO: 983)	TATGAACCACTAAGGCGTcaacctgtgGGCGTGGTGGGCA GCCATTGGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2556)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACGACCAATG GCGTCGCCACCCGCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4129)	HZAFZ_+10-1087153323 P1P2	HZAFZ	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGTGGTGGATCTCTCAC CGT (SEQ ID NO: 984)	TATGAACCACTAAGGCGTcaacctgtgGGGTGGTGGATC TCACCGTgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2557)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACGAGGTGA GAATCCCACCACCCGCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4130)	HZAFZ_+10-1087137923 P1P2	HZAFZ	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCGCGGGACCGGCCGCCA ACA (SEQ ID NO: 985)	TATGAACCACTAAGGCGTcaacctgtgGTGCGCGGGACCGC CGCCAACAgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2558)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACGTTGGGC GGCGGTCCGCGGGCCCAACAAGGTTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4131)	HAT1_+1727911823-P1P2	HAT1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTGATCGTAGCGTCGG AAA (SEQ ID NO: 986)	TATGAACCACTAAGGCGTcaacctgtgGGTGATCGTAGCG TCGGAAAgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2559)	TCACCACGCAGACTTTAGCGAGCTTAGCTCTTAAACGTTCCGGAG CTAGGATCACCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4132)	HAT1_-1727790123-P1P2	HAT1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTCCGAGCTACGATCAC CCG (SEQ ID NO: 987)	TATGAACCACTAAGGCTcaacctgtgG TCCGAGCTACGA TCGTAGCTCGGAACCAACAAGGTCGCACGCCTTAGTGGTCGA (SEQ ID NO: 2550)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGGTGA (SEQ ID NO: 4133)	H4IT_+_172 779897.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACCGACTGACCGGTAGG GAC (SEQ ID NO: 998)	TATGAACCACTAAGGCTcaacctgtgGACCGACTGAGCG TACCGACTGGCCAATTCGCTGGTGA (SEQ ID NO: 2551)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACGTCCCTAC CGTCAGTCCGGTTCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4134)	HDAC1_-_32767752.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAACCGACCGACTGAC GGT (SEQ ID NO: 999)	TATGAACCACTAAGGCTcaacctgtgGGACCGACTGAGCG TCGACGGTCCGGTCCCGAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 2552)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACCACCGTCA GTCGGTCCGGTCCCGAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4135)	HDAC1_-_32767752.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAACGGGAAGGCGAGC AAGA (SEQ ID NO: 990)	TATGAACCACTAAGGCTcaacctgtgCGGACCGGAAGGCG TGCCTCCCGTCCGCCAACAAGGTCGCACGCCTTAGTCGGTTCGATA (SEQ ID NO: 2553)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACTCTGCTC TGCCTCCCGTCCGCCAACAAGGTCGCACGCCTTAGTCGGTTCGATA (SEQ ID NO: 4136)	HDAC1_-_32767773.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTCGGTTACCACCCGG CAG (SEQ ID NO: 991)	TATGAACCACTAAGGCTcaacctgtgGGTCGGTTACGCAC GTTGGTTACCGAGCCCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 2554)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACTTCCTCCC TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACTTCCTCCC (SEQ ID NO: 4137)	HDAC2_-_14292259.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCTCCACGGGAGAG GAA (SEQ ID NO: 992)	TATGAACCACTAAGGCTcaacctgtgGGAGCCTCCACGGG GGAGGAAgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2555)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACTTCCTCCC CCGTGGAGGCCCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4138)	HDAC2_-_14292382.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGCCCCCCTCCTCGCG AGT (SEQ ID NO: 993)	TATGAACCACTAAGGCTcaacctgtgCGTGGCCCCCTCCT CCACGAGTgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2556)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACCACTCGCG AGGAGTGGGGCCCACCAACAAGGTCGCACGCCTTAGTGGTCGAT A (SEQ ID NO: 4139)	HDAC2_+_14292294.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAACCCCAAGTTAGCTCT CA (SEQ ID NO: 994)	TATGAACCACTAAGGCTcaacctgtgGAACCCCAAGTTA CCTTCTGGGGTCCCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 2557)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACTGAGAGC TAACTTGGGGTCCCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4140)	HELQ_-_84376978.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAGCCCTGAGAAGCTAA CTT (SEQ ID NO: 995)	TATGAACCACTAAGGCTcaacctgtgGGAAGCCCTGAGAAG CTAACTTgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2558)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACAAGTTAGC TCTCAGGGCCTCCCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4141)	HELQ_+_84378961.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTCGTGGCCCGGGAGAT GAG (SEQ ID NO: 996)	TATGAAGCACTAAGGCTcaacctgtgCGCGTGGCCCCGG AGATGAGgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2559)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCATCTC CGAGGGCACGAGCCCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4142)	HELQ_+_84376726.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGAACATGGCGGGAGCGG CAG (SEQ ID NO: 997)	TATGAACCACTAAGGCTcaacctgtgGCAGAACATGGGCGG AGCGGCAGgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2560)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTTAAACCTGCCGT CCGGCCATGTCTCTGCCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4143)	HIRA_+_194351943.23-P1	P1	G22orf39_+_19438194.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGCGCGGTGGCCCAG ACA (SEQ ID NO: 998)	TATGAACCACTAAGGCTcaacctgtgGCGCGGCGGTCGC CCCAGACAgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2561)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACTGTCTGG GCCACCCGCGGCGCGCAACAAGGTCGCACGCCTTAGTGGTCGAT A (SEQ ID NO: 4144)	HIRA_+_194352207.23-P1	P1	G22orf39_+_19438207.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCCTCCCGGAACGCCC CGG (SEQ ID NO: 999)	TATGAAGCACTAAGGCTcaacctgtgCTCCCTCCGGGAAG ACCCCCGGgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2572)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACCCCGGGGT CTTCCGGGAGGGGACCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4145)	HIRA_-_19435159.23-P1	P1	G22orf39_-_19435159.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGAGGGGTGCGACGGG GATG (SEQ ID NO: 1000)	TATGAACCACTAAGGCTcaacctgtgGAGGAGgggatgggg gatggttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2573)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACGCATCCCCT GGGGTCCGCTCGTCCGAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4146)	HIRA_+_194392123.23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGGCAACGCCGGGAAG TCA (SEQ ID NO: 1001)	TATGAACCACTAAGGCTcaacctgtgGCCGGGCAACGCC GAAGTCAgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2574)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACTGACTTCC TGCCGTTGCCCGCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4147)	HIRA_-_19418288.23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACCGGGGAGGCCGCCTC AGG (SEQ ID NO: 1002)	TATGAAGCACTAAGGCTcaacctgtgCGGCCCGGGAGCCG CCTGAGGgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2575)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACCCTCAGG CGGGTCCCGGGGCCCAACAAGGTCGCACGCCTTAGTGGTCGAT A (SEQ ID NO: 4148)	HIRA_+_194392231.23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGAACGTCTGACTGGAC TCG (SEQ ID NO: 1003)	TATGAACCACTAAGGCTcaacctgtgCGAACGTCTGACT CGAACTCGgttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2576)	TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACGCGACGTCC AGTCGACGGTCCGGTCCAACAAGGTCGCACGCCTTAGTGGTCGATA (SEQ ID NO: 4149)	HLTF_+_146804273.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGAGCGGCGTCGACGTTCT GAC (SEQ ID NO: 1004)	TATGAACCACTAAGGCGTcaacctgtgGGACGCGCGTCGAC GTCGACGCGCGTCCAACAAGCTCTGCTGGTGA (SEQ ID NO: 2577)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4150)	-HLTF +_148 80:49891 23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCAGGAGGCGCAGGACT GAA (SEQ ID NO: 1005)	TATGAACCACTAAGGCGTcaacctgtgGTCAGGAGGCGCAC GACGTGAAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2578)	TCACCAGCACAGACTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4151)	-HLTF -_148604295 23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CGCACGCCAGCGCGGGCTCC TAT (SEQ ID NO: 1006)	TATGAACCACTAAGGCGTcaacctgtgCGCACGCCAGCGCTC CTGCTATgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2579)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4152)	-HMGB1 -_37040100.2 3-P1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGTGCCGCACGTTCCCC CCA (SEQ ID NO: 1007)	TATGAACCACTAAGGCGTcaacctgtgGCGCGTTGCGCACGT TAAGAAGTCTGCTGGTGA (SEQ ID NO: 2580)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4153)	-HMGB1 -_31039808.2 3-P1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAGCCAATAGGAGCGCG CGC (SEQ ID NO: 1008)	TATGAACCACTAAGGCGTcaacctgtgGTAGCCAATAGGA CGCGCGCGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2581)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4154)	-HMGB1 +_3 1040092.23-P1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACCACGCTCGCGCGGA DCA (SEQ ID NO: 1009)	TATGAACCACTAAGGCGTcaacctgtgGACCACGCTCGCG CGGGACCAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2582)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4155)	-HMGB1 -_31119581.2 3-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACGTCCGCCGCGTCAC CCC (SEQ ID NO: 1010)	TATGAACCACTAAGGCGTcaacctgtgGCACGTCCGCCGCGTG TCACCCCgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2583)	TCACCAGCACAGACTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4156)	-HMGB1 -_31119581.2 3-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGACCCGGCGGAGGTGG GAG (SEQ ID NO: 1011)	TATGAACCACTAAGGCGTcaacctgtgGTGACCCGGCGGAG GTGGGAGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2584)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4157)	-HMGB1 +_3 119161623-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGAGAAGGGACGAAGGTA GGC (SEQ ID NO: 1012)	TATGAACCACTAAGGCGTcaacctgtgGAGAGAAGGGACG AAGGTAGCgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2585)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4158)	-HNRNPA1 -_54674607.2 3-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGCGGTGGTAAAAAAG AGA (SEQ ID NO: 1013)	TATGAACCACTAAGGCGTcaacctgtgGAGGGCGGTGGTAAA AAAGAGAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2586)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4159)	-HNRNPA1 -_54674463.2 3-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTAAAAAAGAGAGAGGG CGA (SEQ ID NO: 1014)	TATGAACCACTAAGGCGTcaacctgtgGGGTAAAAAAGAG AGGGGCGAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2587)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4160)	-HNRNPA1 -_54674459.2 3-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGAGGGTCGAGACCGA GGG (SEQ ID NO: 1015)	TATGAACCACTAAGGCGTcaacctgtgGAGGAGGGTCGAGA CCGAGGGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2588)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4161)	-HNRNPUL1_1 41770913.2 3-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTTCGCCTCCTGACAG GAA (SEQ ID NO: 1016)	TATGAACCACTAAGGCGTcaacctgtgGCTTCGCCTCCTGA CAGGAAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2589)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4162)	-HNRNPUL1_1 41770343.2 3-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCGCCTTCGGCCTCGTG CCT (SEQ ID NO: 1017)	TATGAACCACTAAGGCGTcaacctgtgGTCCGCCTTCGGC CTCGTGCCTgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2590)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4163)	-HNRNPUL1_1 41770338.2 3-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGCGGTGGTGTGGGAAC CAA (SEQ ID NO: 1018)	TATGAACCACTAAGGCGTcaacctgtgGCCCGCGGTGGTGT GGGAACCAAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2591)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4164)	-HNRNPUL2_+_52494886.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCAGGTTCCCAAGCGG AGG (SEQ ID NO: 1019)	TATGAACCACTAAGGCGTcaacctgtgGCGCAGGTTCCCA AGCGGAGGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2592)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4165)	-HNRNPUL2_+_52494668.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTTCGGTCGAACGTCTTCG CAG (SEQ ID NO: 1020)	TATGAACCACTAAGGCGTcaacctgtgGCTTCGGTCGAACG TCTTCGCAGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2593)	TCACCAGCACAGACTTAAGCGTCAACCTGTGGACGCTTAAGCTTCAGTCGTTCATA (SEQ ID NO: 4166)	-HNRNPUL2_+_52494684.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GAGAGCTCGGCGGCTGG GAC (SEQ ID NO: 1021)	TATGAAGCACTAAGGCGTcaacctgtgGaagagtcggagatgg (SEQ ID NO: 2594)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAAGGTCCCAAG CCCGCGGAGCTCTCCAACAAGGTGAGTGCTTAGTGGTTCATA (SEQ ID NO: 4167)	HPf1BP3_+_2 11137783 23-P1P2	HPf1BP3	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGTGGCGGAGGCGA CTAG (SEQ ID NO: 1022)	TATGAACCACTAAGGCGTcaacctgtgGaggtggcggaggcgacc (SEQ ID NO: 2595)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCTAGTGG CCTCGCGCCACCGCGCAACAAGGTGGACTCCTTAGTGGTTCATA (SEQ ID NO: 4168)	HPf1BP3_+_2 11137749 23-P1P2	HPf1BP3	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTCGGCGGCTGGCAG TGG (SEQ ID NO: 1023)	TATGAACCACTAAGGCGTcaacctgtgGgctcggcggctgaactg (SEQ ID NO: 2596)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCCAGTCCC CAGCCGCGGAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4169)	HPf1BP3_+_2 11137775 23-P1P2	HPf1BP3	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGCCCGGTCCACGGA TCT (SEQ ID NO: 1024)	TATGAACCACTAAGGCGTcaacctgtgCAGGCCCGGTCCTCT	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCAGATCGGT GGCAGTGGGGCCTGCCAACAAGGTGGCAGCCTTAGTGGTTCATA (SEQ ID NO: 4170)	HUS1_-_480193022-P1P2	HUS1	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCATCAAGGCGGCTTT CAA (SEQ ID NO: 1025)	TATGAACCACTAAGGCGTcaacctgtgGGCATCAAGGCTTT	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACTTGAAAGCC CGCCTTCATCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4171)	HUS1_+_48319843 23-P1P2	HUS1	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTACCCACACAGAGGCCCG CCG (SEQ ID NO: 1026)	TATGAACCACTAAGGCGTcaacctgtgGTACCCACACAGAGG (SEQ ID NO: 2599)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCCGGCGGG CCCTCTGTGGGTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4172)	HUS1_+_48319323-P1P2	HUS1	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTTTCGCGCCAAAGATC AC (SEQ ID NO: 1027)	TATGAAGCACTAAGGCGTcaacctgtgGAGTTTCGCGCCAA	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAAGTGATCTCT GAGATCACGTTAAGGTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4173)	HUS1B_+_85699123-P1P2	HUS1B	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGCGGGAATCCCCCCT CAG (SEQ ID NO: 1028)	TATGAACCACTAAGGCGTcaacctgtgGCAGGCGGGAATCCC	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCTGAAGGG CCCTCAGGTTCCGCGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4174)	HUS1B_+_85699703-P1P2	HUS1B	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAAGCGTGTGGTGCGAG GTG (SEQ ID NO: 1029)	TATGAACCACTAAGGCGTcaacctgtgGCAAGCGTGTGGTGA	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCACCGTCG CACCACACGCCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4175)	HUS1B_+_85676723-P1P2	HUS1B	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGAGCTGCAGTAGCGCC GGG (SEQ ID NO: 1030)	TATGAAGCACTAAGGCGTcaacctgtgGCTGAGCTGCAGTAG	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCCGGCGGC TACTGCAGTTCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4176)	HUWE1_+_53710912 23-P1P2	HUWE1	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCCGGTATAGAGGTTGT CGC (SEQ ID NO: 1031)	TATGAACCACTAAGGCGTcaacctgtgGGTCCGGTATAGAGG	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTTAAAGGCGAGAA CCTCTACCGTAAGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4177)	HUWE1_+_53709863 23-P1P2	HUWE1	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCTCTCGTGGGATTCGTT CGG (SEQ ID NO: 1032)	TATGAAGCACTAAGGCGTcaacctgtgGTCCTCGTGGGGATT	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCCGAAGAAT TCGCAAGGACGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4178)	HUWE1_+_53710926 23-P1P2	HUWE1	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGAGGCCGCGATATTG GAG (SEQ ID NO: 1033)	TATGAAGCACTAAGGCGTcaacctgtgGCCGGAGGCCGCGCCAT	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAAGCTGCAATAA TGGCGCCTCCGCGCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4179)	INO3_-_120559375 23-P1P2	INO3	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGCGGGTGCTGCTAG CGG (SEQ ID NO: 1034)	TATGAACCACTAAGGCGTcaacctgtgGGAGCGGGTGCTGCTG	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGCGCTAG CAGCACCCGGCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4180)	INO3_-_120554859 23-P1P2	INO3	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGCGGCCACAAGACGCC CGAG (SEQ ID NO: 1035)	TATGAACCACTAAGGCGTcaacctgtgGGGGCGGCCACAAGA	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTGTTAAACCTCGCGGC gaagcttaggagccaagcgCGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2608)	INO3_-_120591040 23-P1P2	INO3	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAATGGCTGCTCTTAAA GG (SEQ ID NO: 1036)	TATGAAGCACTAAGGCGTcaacctgtgGCAATGGCTGCTCTT	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCCTTTAAG AAGCAGCGATTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4182)	INIP_-_115483222 23-P1P2	INIP	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTAGTTTACCTGGGAC CCG (SEQ ID NO: 1037)	TATGAACCACTAAGGCGTcaacctgtgGCGTAGTTTACCTGG	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGGGTGTCC CACGGTAAGTACGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4183)	INIP_-_115480278 23-P1P2	INIP	Pf1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

G93CAACGAGGAACGCTCAAC TCG (SEQ ID NO: 1038)	TATGAAACCACTAAG3GCTcaacctgtgGCGAAGAG3AGCT CAACTCTGgttaagagctaagtgtGCTAAAGTCTTGCTG3TGA (SEQ ID NO: 2611)	TCACCAGCACAGACTTTAGCGACGTTAGCTCTTAAACCGAGTTG ACGTCTCCTCTTGCCCAACAAG3TG3ACGCCTTAGTGCTTCATA (SEQ ID NO: 4184)																			
GAGCCGCGGTTCGCTCT CTG (SEQ ID NO: 1039)	TATGAACCACTAAGGCGCTcaacctgtgGAGCGCGGTTCG GTCTCTGgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2612)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACGAGAAG CGAACCCGCTGCTCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4185)	NO80_-_41408284.2 3-P1P2	NO80	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGAGCCGGAGCAGCCGA CAC (SEQ ID NO: 1040)	TATGAACCACTAAGCGCTcaacctgtgGAGCGAGTCCGAGAGCA GGTGCAGTAGgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2613)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACGTTGTCCCT GCTCCGCTCTCGTCAACGAAGGTTGAACGCCTTAGTGGTTCATA (SEQ ID NO: 4186)	NO80_-_41408445.2 3-P1P2	NO80	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGAACCGCGGTGCTCCAC CGT (SEQ ID NO: 1041)	TATGAACCACTAAGGAGCTcaacctgtgCGCGAACGCGCTGCT CCAGCGTgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2614)	TCACCAGCAGAGACTTTAGCGACGTTAGCTCTTAAACACGGTGG AGCGCGCGGTTTCGCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4187)	NO80_+_41408253.23-P1P2	NO80	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGCACCTCTGGGGCTA TGG (SEQ ID NO: 1042)	TATGAACCACTAAGGCGCTcaacctgtgGAGCACCTCTGGGG CCTATGGgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2615)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACCCATAGCC CCACGAGGTGCTCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4188)	NO80B_-_74682269.2 3-P1P2	NO80B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGACTGTGGAAGCACCT CTG (SEQ ID NO: 1043)	TATGAACCACTAAGGCGCTcaacctgtgGCGGCGTCGGGAAGC ACGTCCTgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2616)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACCAGAGGT GCTCCCACGGCGGCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4189)	NO80B_-_74682279.2 3-P1P2	NO80B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGTAAAGCTCTGCGCGG CGT (SEQ ID NO: 1044)	TATGAACCACTAAGGCGCTcaacctgtgCGAAGTAAAGCTCTG CGCGCCGTgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2617)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACACGCGCGC CACACGCTTACTCCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4190)	NO80B_-_74682268.2 3-P1P2	NO80B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAAGACTCGGAAGTGCGA TGG (SEQ ID NO: 1045)	TATGAACCACTAAGGAGCTcaacctgtgGAAGACTCCGGAAGT CCGGATGGgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2618)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACCCATCGGA CTTCGGAGTCTTCCAACAAGGTTGAACGCCTTAGTGGTTCATA (SEQ ID NO: 4191)	NO80C_+_3077831.23-P1P2	NO80C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGTTCCAAGGGCCGGC GCT (SEQ ID NO: 1046)	TATGAACCACTAAGGCGCTcaacctgtgGAAGTTCCAAGGGC CCCGCGCTgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2619)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACGAGCGCGG GCCTTGGAACTTCCCAACAAGGTTGAACGCCTTAGTGGTTCATA (SEQ ID NO: 4192)	NO80C_+_3077875.23-P1P2	NO80C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGGCGCTGGGAAAAAA GGT (SEQ ID NO: 1047)	TATGAAGCACTAAGGCGCTcaacctgtgCGCCCGCGCGCTGGGA AAAAAGGTgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2620)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACACCCTTTT CCCAGGCGCGGGCCAACAAGGTTGAACGCCTTAGTGGTTCATA (SEQ ID NO: 4193)	NO80C_+_3077884.23-P1P2	NO80C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCCTACAGCACCACCACT AC (SEQ ID NO: 1048)	TATGAAGCACTAAGGAGCTcaacctgtgGCCTCCTACAGCAC GACTACgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2621)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACGGTAGTGGG TGCTGTAGGACGGCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4194)	NO80D_-_206950834.23-P1P2	NO80D	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAGGCGCTCGGCAATTGG AACT (SEQ ID NO: 1049)	TATGAAGCACTAAGGCGCTcaacctgtgGGAGGCGCTCGGCA ATGGCACTgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2622)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACGAGTCGATT GCCAGCGCGCTCCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4195)	NO80D_-_206950711.23-P1P2	NO80D	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
G93CCGGCGCTGTACGAAG AAG (SEQ ID NO: 1050)	TATGAAGCACTAAGGCGCTcaacctgtgCGCCGCGCGCTGTA ACAAGCCGgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2623)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACCTTCTTGT ACAAGCCCCGCGCCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4196)	NO80D_-_206950666.23-P1P2	NO80D	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAGCGCGGAAGCGCGACGAC TTCT (SEQ ID NO: 1051)	TATGAACCACTAAGGCGCTcaacctgtgGTAGCGCGAAGCGCG ACGACTTCgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2624)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACGAAGTTCT GCGCTCCCGCTAGCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4197)	NO80E_-_300076112 3-P1P2	NO80E	P1P2	na	na	HIRIP3_-_300376112 3-P1P2	HIRIP3	na	na	na	na	na	na	na	na	na	na	na
CGGCGGGCCTTGCTGCTGTG AGG (SEQ ID NO: 1052)	TATGAACCACTAAGGCGCTcaacctgtgCGGCGGGCCTTGCT GCTGAGGgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2625)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACCCTCAGCGA GCAAGGCGCCCGCCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4198)	NO80E_+_3000680423-P1P2	NO80E	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCACCGCTAGCTTCTTG CG (SEQ ID NO: 1053)	TATGAAGCACTAAGGCGCTcaacctgtgGTCACCGCTAGCTT CTTCCGgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2626)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACCCGAAGA AGCTACGGGTTGAACCAACAAGGTTGAACGCCTTAGTTGCTTCATA (SEQ ID NO: 4199)	NO80E_+_3000687623-P1P2	NO80E	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CAATAAGCGCTATGGGA G9A (SEQ ID NO: 1054)	TATGAACCACTAAGGCGCTcaacctgtgGAATAAGCGCTATG CGAG9AgttaagagctaagtgtGCTAAAGTCTTGCTGGTGA (SEQ ID NO: 2627)	TCACCAGCAGACGACTTTAGCGACGTTAGCTCTTAAACGTCGTCGCAA TACGCGCTTATTCACAACAAGGTTGACCGCCTTAGTTGGTTCATA (SEQ ID NO: 4200)	NT83_+_153706591.23-P1P2	NT83	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GAGGGTGCCATAGTCGAA G6G (SEQ ID NO: 1055)	TATGAACCACTAAGGCGTcaacctgtgGAGGTCGCCATAGT GTAATGGCCTCGTCCAACAAGGTGTGAGCTCTTAGTTGTTA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4201)	(NTS3_+_15 3700638.23-P1P2	NTS3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTGAGGCAATAAGCGC TAT (SEQ ID NO: 1056)	TATGAACCACTAAGGCGTcaacctgtgGGTGAGGCAATAAG GCGGTATgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4202)	(NTS3_+_15 3700658.23-P1P2	NTS3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGACCGCGGGGGTAC GAC6 (SEQ ID NO: 1057)	TATGAACCACTAAGGCGTcaacctgtgACGGGAGGGCGCGG GTACGACGgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4203)	(NTS6_+_62 028860.23-P1P2	NTS6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGTCGCTTCGGCAACC C6G (SEQ ID NO: 1058)	TATGAACCACTAAGGCGTcaacctgtgCGAGTTCCCTTCGGC ACCCCCGGgttaagagtaagtgtGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4204)	(NTS6_-_52 520277.35.2 3-P1P2	NTS6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGTTACGACGGGAGC AGT (SEQ ID NO: 1059)	TATGAACCACTAAGGCGTcaacctgtgGGGGTTACGACGGGG TCAGCACTgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4205)	(NTS6_+_52 028861.23-P1P2	NTS6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGAGAGCAGAAGCGG TAA (SEQ ID NO: 1060)	TATGAACCACTAAGGCGTcaacctgtgGCGGAGAGCAGAAG GCGGTAAgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4206)	(PMK_-_60027894.2 3-P1P2	PMK	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGCTCGGTTCTGTCT ACC (SEQ ID NO: 1061)	TATGAACCACTAAGGCGTcaacctgtgCGCGGCTCGGTTCT GTCTACGgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4207)	(PMK_+_600 27926.23-P1P2	PMK	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTCGCCGTCAGGGCC CCA (SEQ ID NO: 1062)	TATGAACCACTAAGGCGTcaacctgtgGCGTCGCCGTCAG GCGCCCAgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4208)	(PMK_+_600 27676.23-P1P2	PMK	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCAGCCCTAGGGCCG CAT (SEQ ID NO: 1063)	TATGAACCACTAAGGCGTcaacctgtgGCCCAGCCCTAGG GCCCACATgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTTACCGCT (SEQ ID NO: 4209)	(KAT2A_-_40273376.2 3-P1P2	KAT2A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGTCCGCCCCGCAGCCGA GAG (SEQ ID NO: 1064)	TATGAAGCACTAAGGCGTcaacctgtgGCGTCGCCCCGCAG CGGAGAGgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCAGCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTCTCTCCGCC (SEQ ID NO: 4210)	(KAT2A_-_40273345.2 3-P1P2	KAT2A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCTGGGAAGGTTCCG CCA (SEQ ID NO: 1065)	TATGAACCACTAAGGCGTcaacctgtgGCCTGGGAAAGGTTC CGGCCAgttaagagtaagtgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTGGCGGGA (SEQ ID NO: 4211)	(KAT2A_-_40273303.2 3-P1P2	KAT2A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGCGCCCCGAGCCCC CGG (SEQ ID NO: 1066)	TATGAACCACTAAGGCGTcaacctgtgGCAGGCGCCCCGAG cggcttaagagtaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2639)	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACCCCGGGGCG (SEQ ID NO: 4212)	(KAT2B_+_20 0821102.23-P1P2	KAT2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCAGTACCAGACGCCCG AGA (SEQ ID NO: 1067)	TATGAAGCACTAAGGCGTcaacctgtgCCGGACTACCAGACG AGGCCGAGgttaagagtaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2640)	TCAGCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTGTCGGGG (SEQ ID NO: 4213)	(KAT2B_-_20061907.2 3-P1P2	KAT2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGTCGCCGTCGCTCGG G3G (SEQ ID NO: 1068)	TATGAACCACTAAGGCGTcaacctgtgGCTCGTCGCCGTCGC TCCCGGGgttaagagtaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2641)	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACGCCCGGGA (SEQ ID NO: 4214)	(KAT2B_-_20081967.2 3-P1P2	KAT2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCAGGGTCCGGCTACT GGG (SEQ ID NO: 1069)	TATGAACCACTAAGGCGTcaacctgtgGGCCAGGGTCCGGCT ACTGGGgttaagagtaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2642)	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACCCCAGTTA (SEQ ID NO: 4215)	(KAT5_+_664 73860.23-P1P2	KAT5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
G3GCGGGACTAAGGAACC TGA (SEQ ID NO: 1070)	TATGAAGCACTAAGGCGTcaacctgtgCGGGG3GACTAAGG AACCTGAgttaagagtaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2643)	TCAGCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACTCAGCGTTC (SEQ ID NO: 4216)	(KAT5_-_65479364.2 3-P1P2	KAT5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTGAGTCGGGTGGCCCG G3G (SEQ ID NO: 1071)	TATGAACCACTAAGGCGTcaacctgtgGGTGAGTCGGGTGG CCCGG3GgttaagagtaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2644)	TCACCAGCAGAGACTTTACGCAGCTTAGCTGTTAAACCCCCGGGG (SEQ ID NO: 4217)	(KAT5_-_65479771.2 3-P1P2	KAT5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCCTGGCGCAGAGCAGCGGG (SEQ ID NO: 1072)	TATGAACCACTAAGGCGTccaccttgttgGCCTGGCGCAGAGCAGCGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2645)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCTGCTCTGCGCCAGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4218)	1	KAT6A_-_41909290.23-P1P2	KAT6A	P1P2	na	na	na	na	na	na	na	na	na
GGAGATGCCCCAAACTGACA (SEQ ID NO: 1073)	TATGAACCACTAAGGCGTccaccttgttgGGAGATGCCCCAAACTGACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2646)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTCAGTTTGGGGCATCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4219)	1	KAT6A_-_41909485.23-P1P2	KAT6A	P1P2	na	na	na	na	na	na	na	na	na
GGAGCTCACTGTCTCCAAGA (SEQ ID NO: 1074)	TATGAACCACTAAGGCGTccaccttgttgGGAGCTCACTGTCTCCAAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2647)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTTGGAGACAGTGAGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4220)	1	KAT6A_+_41909489.23-P1P2	KAT6A	P1P2	na	na	na	na	na	na	na	na	na
GACAACAGGGGGACACAAAA (SEQ ID NO: 1075)	TATGAACCACTAAGGCGTccaccttgttgGACAACAGGGGGACACAAAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2648)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTTTGTGTCCCCCTGTTGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4221)	1	KAT6B_-_76586407.23-P1P2	KAT6B	P1P2	na	na	na	na	na	na	na	na	na
GACAGGGGGACACAAAATGG (SEQ ID NO: 1076)	TATGAACCACTAAGGCGTccaccttgttgGACAGGGGGACACAAAATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2649)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATTTTGTGTCCCCCTGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4222)	1	KAT6B_-_76586410.23-P1P2	KAT6B	P1P2	na	na	na	na	na	na	na	na	na
GTGTCAACGACAACAACAGG (SEQ ID NO: 1077)	TATGAACCACTAAGGCGTccaccttgttgGTGTCAACGACAACAACAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2650)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTGTTGTTGTCGTTGACACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4223)	1	KAT6B_-_76586396.23-P1P2	KAT6B	P1P2	na	na	na	na	na	na	na	na	na
GACTGATACAGAGGCCGCCA (SEQ ID NO: 1078)	TATGAACCACTAAGGCGTccaccttgttgGACTGATACAGAGGCCGCCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2651)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGCGGCCTCTGTATCAGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4224)	1	KAT7_-_47866121.23-P1P2	KAT7	P1P2	na	na	na	na	na	na	na	na	na
GAGGGATCGTCCGCAGGATT (SEQ ID NO: 1079)	TATGAACCACTAAGGCGTccaccttgttgGAGGGATCGTCCGCAGGATTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2652)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAATCCTGCGGACGATCCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4225)	1	KAT7_-_47866099.23-P1P2	KAT7	P1P2	na	na	na	na	na	na	na	na	na
GTGTATCAGTCCCAATCCTG (SEQ ID NO: 1080)	TATGAACCACTAAGGCGTccaccttgttgGTGTATCAGTCCCAATCCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2653)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGGATTGGGACTGATACACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4226)	1	KAT7_+_47866087.23-P1P2	KAT7	P1P2	na	na	na	na	na	na	na	na	na
GACTTCCCTTCCCGCGATGG (SEQ ID NO: 1081)	TATGAACCACTAAGGCGTccaccttgttgGACTTCCCTTCCCGCGATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2654)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATCGCGGGAAGGGAAGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4227)	1	KAT8_-_31129008.23-P1	KAT8	P1	na	na	na	na	na	na	na	na	na
GGCACAGGGAGCTGCTGCGG (SEQ ID NO: 1082)	TATGAACCACTAAGGCGTccaccttgttgGGCACAGGGAGCTGCTGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2655)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCAGCAGCTCCCTGTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4228)	1	KAT8_-_31129029.23-P1	KAT8	P1	na	na	na	na	na	na	na	na	na
GGTTGCGGCGGGGACTTCAG (SEQ ID NO: 1083)	TATGAACCACTAAGGCGTccaccttgttgGGTTGCGGCGGGGACTTCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2656)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGAAGTCCCCGCCGCAACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4229)	1	KAT8_-_31129050.23-P1	KAT8	P1	na	na	na	na	na	na	na	na	na
GCTGCAGGAGAAGGAGTCCC (SEQ ID NO: 1084)	TATGAACCACTAAGGCGTccaccttgttgGCTGCAGGAGAAGGAGTCCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2657)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGACTCCTTCTCCTGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4230)	1	KAT8_-_31141602.23-P2	KAT8	P2	na	na	na	na	na	na	na	na	na
GGGTTCTGACAGTAAATCTA (SEQ ID NO: 1085)	TATGAACCACTAAGGCGTccaccttgttgGGGTTCTGACAGTAAATCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2658)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGATTTACTGTCAGAACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4231)	1	KAT8_+_31141331.23-P2	KAT8	P2	na	na	na	na	na	na	na	na	na
GTCAGGATGTAAAAGACGAA (SEQ ID NO: 1086)	TATGAACCACTAAGGCGTccaccttgttgGTCAGGATGTAAAAGACGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2659)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCGTCTTTTACATCCTGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4232)	1	KAT8_+_31141409.23-P2	KAT8	P2	na	na	na	na	na	na	na	na	na
GAAGCAGTAGGATCGGCCAG (SEQ ID NO: 1087)	TATGAACCACTAAGGCGTccaccttgttgGAAGCAGTAGGATCGGCCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2660)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGGCCGATCCTACTGCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4233)	1	KDM4A_-_44115875.23-P1P2	KDM4A	P1P2	na	na	na	na	na	na	na	na	na
GCAGATGCCGACTTTAGAGG (SEQ ID NO: 1088)	TATGAACCACTAAGGCGTccaccttgttgGCAGATGCCGACTTTAGAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2661)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTCTAAAGTCGGCATCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4234)	1	KDM4A_-_44115826.23-P1P2	KDM4A	P1P2	na	na	na	na	na	na	na	na	na

GCGCGCTGCTGGAAAAGC AGT (SEQ ID NO: 1089)	TATGAAGCACATAAG3GCGTcaacctgtgGGGCCTGCTGGAA AAGCACGTgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2682)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACACTGCTTT TCCACGCAGGCGCCCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4235)	KCNMA_-44115863.2 3-P1P2	KCNMA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACGCCGAGGGACCCC CTA (SEQ ID NO: 1090)	TATGAACCACTAAGGCGTccaacctgtgCACCCCGAGGGA GCCCGTTAgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2683)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACTAGGGGG TCCGTTCGGGGTGCCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4236)	KEAP1_+_10 8713787.23-P1P2	KEAP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCTTGTCCTCAGGCG GTA (SEQ ID NO: 1091)	TATGAACCACTAAGGCGTccaacctgtgCGTCCGTCGCCTCA GGCGTGTTAgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2684)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACTACCGCCT GAGGCCGAGGGCCCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4237)	KEAP1_-_106135398.2 3-P1P2	KEAP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGACGCCGACGCCCCC CGA (SEQ ID NO: 1092)	TATGAACCACTAAGGCGTccaacctgtgGTGGAGCCGAGGCC GCCCCCGAgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2685)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACTCGGGGC GCCTCCGGCTCCACCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4238)	KEAP1_+_10 8713847.23-P1P2	KEAP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAATGCGCGTGGCGCGAA CTT (SEQ ID NO: 1093)	TATGAACCACTAAGGCGTccaacctgtgGAATGCGCGTGGCGA CGAACTTgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2686)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACAAGTTCGC GCGACGCGCATTCCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4239)	LIG1_+_4867 3564.23-P1P2	LIG1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGAACTTGCGGACTG CAG (SEQ ID NO: 1094)	TATGAACCACTAAGGCGTccaacctgtgGCGCGAACTTGCGG ACTGCAGgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2687)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACCTGCAGT CCCAAGTTCGCGCGCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4240)	LIG1_+_4867 3564.23-P1P2	LIG1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTCTGCGCGCGAATG CCG (SEQ ID NO: 1095)	TATGAACCACTAAGGCGTccaacctgtgCGCGTCGCGCGTG CAATGCCGgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2688)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACCGGCAATT CGCGCGCGCAGACGCCAACAAGGTGGACCGCCTTAGTGGTTCAT A (SEQ ID NO: 4241)	LIG1_+_4867 3678.23-P1P2	LIG1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGGCGCTCCAACCGT CGT (SEQ ID NO: 1096)	TATGAACCACTAAGGCGTccaacctgtgGCAGGGCGCTCCAA CCGTCGGgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2689)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACACGAGCGG ITCGGAGCGCGTGCGCCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4242)	LIG3_-_33307576.2 3-P1P2	LIG3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGCTCCGGGCGCTCAC CGT (SEQ ID NO: 1097)	TATGAACCACTAAGGCGTccaacctgtgGCGGCTCCGGGCGC TCACCGTgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2690)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACACGGTGA GCGCCGCGGAGCGCCGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4243)	LIG3_+_33308 7810.23-P1P2	LIG3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAGCAAGTTCCGCGAGGC CTA (SEQ ID NO: 1098)	TATGAAGCACTAAGGCGTccaacctgtgCGAAGCAAGTTCCG AGGCCTAgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2691)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACTAGGCCT CGAAACTTGCTCCGCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4244)	LIG3_-_33307618.2 3-P1P2	LIG3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTTCAGGTGGTGGAAGGG TGT (SEQ ID NO: 1099)	TATGAACCACTAAGGCGTccaacctgtgCGCGTCAGgttaagagtaag cgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2692)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTTAAACACACCCTT CCCACCTGAACGCCGCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4245)	LIG4_+_1088 67815.23-P1	LIG4	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGGCCAGTTCCGGGCT TAG (SEQ ID NO: 1100)	TATGAACCACTAAGGCGTccaacctgtgGGCGGCCAGCTTCC TGAGTTAGgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2693)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACCTAAGCCC TGAAGCTGGGCGCCCAAGAAAGGTGGACCGCCTTAGTGGTTCAT A (SEQ ID NO: 4246)	LIG4_+_1088 67954.23-P1	LIG4	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTCAGGGTTGAGGTGCA GGT (SEQ ID NO: 1101)	TATGAAGCACTAAGGCGTccaacctgtgCGTTCAGGGCTTGAC GTCAGgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2694)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACACCTGAC GTCAAGCCTCTGAAACCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4247)	LIG4_+_1088 67926.23-P1	LIG4	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATCGCTAGGTCCTTGGA GCA (SEQ ID NO: 1102)	TATGAACCACTAAGGCGTccaacctgtgGATCGCTAGGTCCT CGGAGCAgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2695)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACGTAGCTCA GGACCTACGGGATCCGCTTAAACCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4248)	LIG4_-_108870652 23-P2	LIG4	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCCAAGATTCCCTAGG TCC (SEQ ID NO: 1103)	TATGAACCACTAAGGCGTccaacctgtgAGGGCCAAGATCCC TAGGTCCgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2696)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACGGGATCC TGGGATCTTGGCCCCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4249)	LIG4_-_108870645 23-P2	LIG4	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTTGGAAGGCCAAGATC CCT (SEQ ID NO: 1104)	TATGAACCACTAAGGCGTccaacctgtgGGTTGGAAGGCCAA GATCCCTgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2697)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACACGGGATC TTGGCCTTCCACCCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4250)	LIG4_-_108870639 23-P2	LIG4	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGAGGCCCCGGGTT GAG (SEQ ID NO: 1105)	TATGAACCACTAAGGCGTccaacctgtgGGCCGAGGCCCCGG GTTGAGgttaagagtaagcgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2698)	TCACGACGCAGAGTTTACGGCAGCTTAGCTGTTAAACCTCAACCCC CCGGGCCTCGGCGCCCAACAAGGTGGACCGCCTTAGTGGTTCATA (SEQ ID NO: 4251)	MAD2L2_+_117411173.23-P1P2	MAD2L2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGCCGCGGGGAGCCACCGAC (SEQ ID NO: 1106)	TATGAACCACTAAGGCGTccaccttgttgGgccgcggggAGCCACCGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2679)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCGGTGGCTCCCCGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4252)	1	MAD2L2_-_11741158.23-P1P2	MAD2L2	P1P2	na	na	na	na	na	na	na	na	na
GGCGGGAGGCGCGATCGGTC (SEQ ID NO: 1107)	TATGAACCACTAAGGCGTccaccttgttgGGCGGGAGGCGCGATCGGTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2680)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGACCGATCGCGCCTCCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4253)	1	MAD2L2_+_11741155.23-P1P2	MAD2L2	P1P2	na	na	na	na	na	na	na	na	na
GAGCGCCAGGTACCACGAGT (SEQ ID NO: 1108)	TATGAACCACTAAGGCGTccaccttgttgGAGCGCCAGGTACCACGAGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2681)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACTCGTGGTACCTGGCGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4254)	1	MAU2_+_19431732.23-P1P2	MAU2	P1P2	na	na	na	na	na	na	na	na	na
GGAAGCAGGCGCGCGAGACG (SEQ ID NO: 1109)	TATGAACCACTAAGGCGTccaccttgttgGGAAGCAGGCGCGCGAGACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2682)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTCTCGCGCGCCTGCTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4255)	1	MAU2_+_19431606.23-P1P2	MAU2	P1P2	na	na	na	na	na	na	na	na	na
GTGGCGGCGGCTTGTTGTTG (SEQ ID NO: 1110)	TATGAACCACTAAGGCGTccaccttgttgGTGGCGGCGGCTTGTTGTTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2683)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAACAACAAGCCGCCGCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4256)	1	MAU2_-_19431659.23-P1P2	MAU2	P1P2	na	na	na	na	na	na	na	na	na
GACTCCGCCATAGAGCAGGG (SEQ ID NO: 1111)	TATGAACCACTAAGGCGTccaccttgttgGACTCCGCCATAGAGCAGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2684)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCTGCTCTATGGCGGAGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4257)	1	MBD2_+_51750822.23-P1P2	MBD2	P1P2	na	na	na	na	na	na	na	na	na
GCGGGTCTCCGGGATTCCAA (SEQ ID NO: 1112)	TATGAACCACTAAGGCGTccaccttgttgGcgggTCTCCGGGATTCCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2685)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGGAATCCCGGAGACCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4258)	1	MBD2_+_51750973.23-P1P2	MBD2	P1P2	na	na	na	na	na	na	na	na	na
GGATTCCAAGGGCTCGGTTA (SEQ ID NO: 1113)	TATGAACCACTAAGGCGTccaccttgttgGGATTCCAAGGGCTCGGTTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2686)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAACCGAGCCCTTGGAATCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4259)	1	MBD2_+_51750962.23-P1P2	MBD2	P1P2	na	na	na	na	na	na	na	na	na
GCCGACAGCCCCGACCTTCT (SEQ ID NO: 1114)	TATGAACCACTAAGGCGTccaccttgttgGCCGACAGCCCCGACCTTCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2687)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGAAGGTCGGGGCTGTCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4260)	1	MBD3_-_1592569.23-P1P2	MBD3	P1P2	na	na	na	na	na	na	na	na	na
GCGGCCGCGCAGGGCCTCAA (SEQ ID NO: 1115)	TATGAACCACTAAGGCGTccaccttgttgGCGGCCGCGCAGGGCCTCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2688)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGAGGCCCTGCGCGGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4261)	1	MBD3_+_1592484.23-P1P2	MBD3	P1P2	na	na	na	na	na	na	na	na	na
GGAACAGCCGCGGCAAGTGG (SEQ ID NO: 1116)	TATGAACCACTAAGGCGTccaccttgttgGgAACAGCCGCGGCAAGTgggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2689)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACTTGCCGCGGCTGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4262)	1	MBD3_+_1592727.23-P1P2	MBD3	P1P2	na	na	na	na	na	na	na	na	na
GCCCAGCGCCGCAACGCCCA (SEQ ID NO: 1117)	TATGAACCACTAAGGCGTccaccttgttgGCCCAGCGCCGCAACGCCCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2690)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGGCGTTGCGGCGCTGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4263)	1	MBD4_-_129158724.23-P1P2	MBD4	P1P2	na	na	na	na	na	na	na	na	na
GCCGGTGCCGTGAGCTGAAG (SEQ ID NO: 1118)	TATGAACCACTAAGGCGTccaccttgttgGCCGGTGCCGTGAGCTGAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2691)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTCAGCTCACGGCACCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4264)	1	MBD4_+_129158749.23-P1P2	MBD4	P1P2	na	na	na	na	na	na	na	na	na
GGCCGCAACGCCCAGGGTGT (SEQ ID NO: 1119)	TATGAACCACTAAGGCGTccaccttgttgGGCCGCAACGCCCAGGGTGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2692)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACACCCTGGGCGTTGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4265)	1	MBD4_-_129158730.23-P1P2	MBD4	P1P2	na	na	na	na	na	na	na	na	na
GGGGCGCGATCTGGATTTGG (SEQ ID NO: 1120)	TATGAACCACTAAGGCGTccaccttgttgGGGGCGCGATCTGGATTTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2693)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCAAATCCAGATCGCGCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4266)	2	MCM10_-_13203752.23-P1P2	MCM10	P1P2	na	na	na	na	na	na	na	na	na
GGGGCGTGCGTGTGACCCTC (SEQ ID NO: 1121)	TATGAACCACTAAGGCGTccaccttgttgGGGGCGTGCGTGTGACCCTcgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2694)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGGGTCACACGCACGCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4267)	2	MCM10_-_13203645.23-P1P2	MCM10	P1P2	na	na	na	na	na	na	na	na	na
GGTGAACGAAGAAGGCGTCC (SEQ ID NO: 1122)	TATGAACCACTAAGGCGTccaccttgttgGGTGAACGAAGAAGGCGTCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2695)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGACGCCTTCTTCGTTCACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4268)	2	MCM10_-_13203599.23-P1P2	MCM10	P1P2	na	na	na	na	na	na	na	na	na

GGATCGTGGTACTGCTAT GC (SEQ ID NO: 1123)	TATGAAGCACTAAGGCGTcaacctgtggGGATGGTGGTAGT	TCACCAGCACAGACTTTAGCGCAGCTTAAGCCATAGCA2 GTACCACGATCCCAAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2689)	MCM2_-_127317315 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTAGTGGCGGCAGAGGA TCG (SEQ ID NO: 1124)	TATGAACCACTAAGGCGTcaacctgtgGGTAGTGGTGGAG	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCATCCTC2 TCCGCCACGTACCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2690)	MCM2_-_127317301 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGTTGCTGTAGTGGCG GAG (SEQ ID NO: 1125)	TATGAACCACTAAGGCGTcaacctgtgGTTGTGCTGTAGTG	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCTCCGCC2 AGTACGACGCAACCACCAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2691)	MCM2_-_127317294 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTGGCCGCAGGCGAAT CCT (SEQ ID NO: 1126)	TATGAACCACTAAGGCGTcaacctgtgCCGTCGCCGCAGG	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCATCGAT1 GCCTGCCGCCACGCCAACAAGGTCGACGCCTTAGTGGTTCAT A (SEQ ID NO: 2692)	MCM6_+_13 6633847 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTTGTGGCGGTCGA GCG (SEQ ID NO: 1127)	TATGAACCACTAAGGCGTcaacctgtgGGACTTAGCTGTCGCGA	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCGCTCGA1 CCCGCCACAAGTCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2693)	MCM6_+_13 6633894 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAAATATGGACCTCTG CGG (SEQ ID NO: 1128)	TATGAACCACTAAGGCGTcaacctgtgGGCAAATATGTGGAC	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCCGCGAG1 GTCCATATTTGCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2694)	MCM6_+_13 6633919 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATCGCTGCCGAGGGCC GTG (SEQ ID NO: 1129)	TATGAACCACTAAGGCGTcaacctgtgGATCGCTGCCGAG	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCTCAACG1 GGCCCGTCGATTGGTAGAGCTAAGCGTCGATCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2702)	MCM7_-_99696963 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCGCTTGGCGGGGCT CAG (SEQ ID NO: 1130)	TATGAACCACTAAGGCGTcaacctgtgGCCGCGCTTGGGG	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCTCGTGA1 CGCCAAGCGCGGGGCCAACAAGGTCGACGCCTTAGTGGTTCAT A (SEQ ID NO: 2703)	MCM7_+_99 6698070 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAGGTCCCGGTGAAAT CAG (SEQ ID NO: 1131)	TATGAACCACTAAGGCGTcaacctgtgGGAAGGTCCCGGT	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCTCGATT1 ACCGGGGACCTCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2704)	MCM7_+_99 6698070 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGAAGTAGACGTAGCC TCA (SEQ ID NO: 1132)	TATGAAGCACTAAGGCGTcaacctgtgCCGAAGTAGACCG	TCAGCAGCAGACAGCTTAAGCGCAGCTTAAGCTCGATA1 CGTCTACTTCCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2705)	MCRS1_+_4 9961900 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGGACCGGGGTTAGA ATC (SEQ ID NO: 1133)	TATGAACCACTAAGGCGTcaacctgtgGGCCGGACCGGGT	TCACCAGCAGACAGCTTTAAGCGCAGCTTAGCTCTTAAA1 CACAGAATCGcttaagagctaagcgcCGTAAAGCTCTGCTGGTTCA (SEQ ID NO: 2706)	MCRS1_+_4 9961887 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCGCGCGCGGAGAAAT TGT (SEQ ID NO: 1134)	TATGAACCACTAAGGCGTcaacctgtgGGTCGCGCGGAG	TCACCAGCAGACAGCTTTAAGCGCAGCTTAGCTCTTAAA1 TCGCGCGCGAACCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2707)	MCRS1_+_4 9961949 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGCGCGTCGGACTTAA AGT (SEQ ID NO: 1135)	TATGAAGCACTAAGGCGTcaacctgtgCGCGCGTCGGACT	TCACCAGCAGACAGCTTAAGCGCAGCTTAAGCCTTTAAG1 TCGCCGCGCGACGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2708)	MDC1_+_30 6835020 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTGCGCGTGCAACGGAT CGATAG (SEQ ID NO: 1136)	TATGAACCACTAAGGCGTcaacctgtgGGCTGCGTGCAAC	TCACCAGCAGACAGCTTTAAGCGCAGCTTAGCTCTTAAG1 TTCGCGCGCAGCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2709)	MDC1_-_30686387 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGTTGCGCGCTTGCAG TCG (SEQ ID NO: 1137)	TATGAACCACTAAGGCGTcaacctgtgGGGGGTTGCGCGC	TCACCAGCAGACAGCTTTAAGCGCAGCTTAGCTCTTAAA1 AACGCCGCAACGCCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2683)	MDC1_-_30685595 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAAGTCAGCCAAGTCTCG CCG (SEQ ID NO: 1138)	TATGAAGCACTAAGGCGTcaacctgtgGTGAAGTCAGCCA	TCAGCAGCAGACAGCTTTAAGCGCAGCTTAGCTCTTAAA1 CTTGGCTGGCTTCCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2684)	MDM2_+_69 2201973 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAAGAAGCGCGAGCCC GAG (SEQ ID NO: 1139)	TATGAACCACTAAGGCGTcaacctgtgGGCAAGAAGCGCG	TCACCAGCAGACAGCTTTAAGCGCAGCTTCGGGCGC1 TCGCGCTTCTTGCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 2685)	MDM2_-_89202652 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

G3GCCCTGTGTGTCGGAA AGA (SEQ ID NO: 1140)	TATGAAGCACTAAG3GCGTccaacttgtgG3GCCTGTGTGTC AGACAGAGGCGCCCAACAAG3GTGGACGCCTTAGTGTCGATA (SEQ ID NO: 4286)					MDM2_-_69020028.2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACATGGCGATGCACA ACA (SEQ ID NO: 1141)	TATGAACCACTAAGGCGTccaacttgtgG3AACATGGCGATG CACAACAgtttaagagctaagctgtGCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4287)					MEAF6_+_3 7980328.23-P1P2	MEAF6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGGCGTGTCCGGCAT CTG (SEQ ID NO: 1142)	TATGAACCACTAAGGCGTccaacttgtgG3CCGGCGTGTCCG GGACACCCGCGCCACAAGGTGGACGCGTTAGTGGTTCATA (SEQ ID NO: 4288)					MEAF6_-_37980319.2 3-P1P2	MEAF6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGTGTCCGGGCATCTG CGG (SEQ ID NO: 1143)	TATGAACCACTAAGGCGTccaacttgtgCGGCGTGTCCGGGA TCCCGGACACCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4289)					MEAF6_-_37980322.2 3-P1P2	MEAF6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GATAACACGGGACGCGGG CGG (SEQ ID NO: 1144)	TATGAACCACTAAGGCGTccaacttgtgGATAACACGGGAG CTTCCTGTTTATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4290)					MGEA5_+_1 03578175.23-P1P2	MGEA5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAGCAGAGGGAGAGC TCG (SEQ ID NO: 1145)	TATGAACCACTAAGGCGTccaacttgtgGGCAGCAGAGGGA GAGCTCGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4291)					MGEA5_+_1 03578146.23-P1P2	MGEA5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
G3GCGCCCAGCTTAAAG CCG (SEQ ID NO: 1146)	TATGAACCACTAAGGCGTccaacttgtgCGCGCCCCAGCTT AAAGCCGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4292)					MGEA5_+_1 03577995.23-P1P2	MGEA5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGACGCCCGGCAGGTCC TCG (SEQ ID NO: 1147)	TATGAACCACTAAGGCGTccaacttgtgCGCGACGCCCGGCAG GTCCTCGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4293)					MGMT_-_13126556.39 23-P1P2	MGMT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAAGGCTGGGGCAACA CCT (SEQ ID NO: 1148)	TATGAACCACTAAGGCGTccaacttgtgGGAAAGGCTGGGGC AACACCTgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4294)					MGMT_+_13 1295722.23-P1P2	MGMT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCGCAACCGCGAGGAC CTG (SEQ ID NO: 1149)	TATGAAGCACTAAGGCGTccaacttgtgCGTGCGCAACCGCG AGGACCTGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4295)					MGMT_+_13 1286524.23-P1P2	MGMT	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGCGCGCTGGAGCAGAG CAG (SEQ ID NO: 1150)	TATGAACCACTAAGGCGTccaacttgtgCGCGCGCGCTGGAG CAGAGCAGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4296)					MLH1_-_37035083.2 3-P1P2	MLH1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTGGCAGCGGCGTTATT CGG (SEQ ID NO: 1151)	TATGAACCACTAAGGCGTccaacttgtgGCGTGGCAGCGGT TATTCGGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4297)					MLH1_-_37035097.2 3-P1P2	MLH1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
G3GGCAGCGTCACGCGAACC CGG (SEQ ID NO: 1152)	TATGAACCACTAAGGCGTccaacttgtgG3GCAGCGTCACG CGAACCCGGgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4298)					MLH1_+_3703280.23-P1P2	MLH1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTGCGAAGGCGGTTGG TGT (SEQ ID NO: 1153)	TATGAACCACTAAGGCGTccaacttgtgGCGTGCGAAGGCG GTTGGTGTgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4299)					MLH3_+_765 18106.23-P1P2	MLH3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCAGCGAGCCTCAAGAT CCA (SEQ ID NO: 1154)	TATGAACCACTAAGGCGTccaacttgtgGGGCAGCGAGCCTCA AGATCCAgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4300)					MLH3_+_765 18140.23-P1P2	MLH3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCTCCGGACGACCAACCG CCT (SEQ ID NO: 1155)	TATGAACCACTAAGGCGTccaacttgtgGTCTCCGGACGACC AACCGCCTgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4301)					MLH3_-_75018124.2 3-P1P2	MLH3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACGCGAAACTAGAGAC CGT (SEQ ID NO: 1156)	TATGAACCACTAAGGCGTccaacttgtgGAACGCGAAACTAG ACCGTgttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4302)					MMS19_-_99258162.2 3-P1P2	MMS19	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na

GGAGGGAGCGGCCTCAGTAC (SEQ ID NO: 1157)	TATGAACCACTAAGGCGTccaccttgttgGGAGGGAGCGGCCTCAGTACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2730)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTACTGAGGCCGCTCCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4303)	1	MMS19_+_99257963.23-P1P2	MMS19	P1P2	na	na	na	na	na	na	na	na	na
GTCGTGCACGAGGCCCCATA (SEQ ID NO: 1158)	TATGAACCACTAAGGCGTccaccttgttgGTCGTGCACGAGGCCCCATAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2731)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTATGGGGCCTCGTGCACGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4304)	1	MMS19_-_99258097.23-P1P2	MMS19	P1P2	na	na	na	na	na	na	na	na	na
GAAATGTGAGCGGGTACGAC (SEQ ID NO: 1159)	TATGAACCACTAAGGCGTccaccttgttgGAAATGTGAGCGGGTACGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2732)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCGTACCCGCTCACATTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4305)	2	MMS22L_+_97730961.23-P1P2	MMS22L	P1P2	na	na	na	na	na	na	na	na	na
GAGCCGGGTACGCGGGATAC (SEQ ID NO: 1160)	TATGAACCACTAAGGCGTccaccttgttgGAGCCGGGTACGCGGGATACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2733)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTATCCCGCGTACCCGGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4306)	2	MMS22L_+_97730904.23-P1P2	MMS22L	P1P2	na	na	na	na	na	na	na	na	na
GGAACGCTCCCTCCCCGGCT (SEQ ID NO: 1161)	TATGAACCACTAAGGCGTccaccttgttgGGAACGCTCCCTCCCCGGCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2734)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGCCGGGGAGGGAGCGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4307)	2	MMS22L_-_97730796.23-P1P2	MMS22L	P1P2	na	na	na	na	na	na	na	na	na
GCTGCCAAGCGCCTGTTGGT (SEQ ID NO: 1162)	TATGAACCACTAAGGCGTccaccttgttgGCTGCCAAGCGCCTGTTGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2735)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCAACAGGCGCTTGGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4308)	2	MNAT1_-_61201494.23-P1P2	MNAT1	P1P2	na	na	na	na	na	na	na	na	na
GGGCGCCTGCGAGAGTCTGT (SEQ ID NO: 1163)	TATGAACCACTAAGGCGTccaccttgttgGGGCGCCTGCGAGAGTCTGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2736)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACAGACTCTCGCAGGCGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4309)	2	MNAT1_-_61201567.23-P1P2	MNAT1	P1P2	na	na	na	na	na	na	na	na	na
GTCCCTTCGCGAAAGAGCAC (SEQ ID NO: 1164)	TATGAACCACTAAGGCGTccaccttgttgGTCCCTTCGCGAAAGAGCACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2737)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTGCTCTTTCGCGAAGGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4310)	2	MNAT1_+_61201445.23-P1P2	MNAT1	P1P2	na	na	na	na	na	na	na	na	na
GCCGGCCCAGGATGTAGAGC (SEQ ID NO: 1165)	TATGAACCACTAAGGCGTccaccttgttgGCCGGCCCAGGATGTAGAGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2738)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCTCTACATCCTGGGCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4311)	1	MORF4L1_-_79165256.23-P1P2	MORF4L1	P1P2	na	na	na	na	na	na	na	na	na
GCTAACGGCGCAGGAGATAG (SEQ ID NO: 1166)	TATGAACCACTAAGGCGTccaccttgttgGCTAACGGCGCAGGAGATAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2739)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTATCTCCTGCGCCGTTAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4312)	1	MORF4L1_-_79165449.23-P1P2	MORF4L1	P1P2	na	na	na	na	na	na	na	na	na
GCTGCGGGCGCTGGCAAATC (SEQ ID NO: 1167)	TATGAACCACTAAGGCGTccaccttgttgGCTGCGGGCGCTGGCAAATCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2740)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGATTTGCCAGCGCCCGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4313)	1	MORF4L1_-_79165238.23-P1P2	MORF4L1	P1P2	na	na	na	na	na	na	na	na	na
GCGACCACGGAAGCACCCCT (SEQ ID NO: 1168)	TATGAACCACTAAGGCGTccaccttgttgGCGACCACGGAAGCACCCCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2741)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGGGGTGCTTCCGTGGTCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4314)	1	MPG_+_128226.23-P1P2	MPG	P1P2	na	na	na	na	na	na	na	na	na
GCGCTCCGCCCCGGTCCTAG (SEQ ID NO: 1169)	TATGAACCACTAAGGCGTccaccttgttgGCGCTCCGCCCCGGTCCTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2742)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTAGGACCGGGGCGGAGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4315)	1	MPG_-_128233.23-P1P2	MPG	P1P2	na	na	na	na	na	na	na	na	na
GTGGGACTCGGACCCCGGCG (SEQ ID NO: 1170)	TATGAACCACTAAGGCGTccaccttgttgGTGGGACTCGGACCCCGGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2743)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCGGGGTCCGAGTCCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4316)	1	MPG_+_128262.23-P1P2	MPG	P1P2	na	na	na	na	na	na	na	na	na
GAGTCTGTCGGTCAGACCAA (SEQ ID NO: 1171)	TATGAACCACTAAGGCGTccaccttgttgGAGTCTGTCGGTCAGACCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2744)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGGTCTGACCGACAGACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4317)	1	MPLKIP_+_40174449.23-P1P2	MPLKIP	P1P2	na	na	na	na	na	na	na	na	na
GCAGTTCCGGTGGGAGAACG (SEQ ID NO: 1172)	TATGAACCACTAAGGCGTccaccttgttgGCAGTTCCGGTGGGAGAACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2745)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTTCTCCCACCGGAACTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4318)	1	MPLKIP_+_40174195.23-P1P2	MPLKIP	P1P2	na	na	na	na	na	na	na	na	na
GCTCCACCCGGACCAGGGTA (SEQ ID NO: 1173)	TATGAACCACTAAGGCGTccaccttgttgGCTCCACCCGGACCAGGGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2746)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACCCTGGTCCGGGTGGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4319)	1	MPLKIP_-_40174132.23-P1P2	MPLKIP	P1P2	na	na	na	na	na	na	na	na	na

GGGCGTAAACCTGAATTCCG (SEQ ID NO: 1174)	TATGAAGCACATAAGGCGTcaacctgtgGGCGTAAACCTGAATTCCGcgttaagtgcctaagtgcctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 2747)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACCCGGAATTCG ACGTTTACGGGCCCAACGAGGTGGAGCTCCTTAGTGGTTCATA (SEQ ID NO: 4320)	MRE11A_94226894.2 3-P1P2	MRE11A	P1P2	ANKRD49_-94226894.2 3-P1P2	ANKRD49	P1P2	na	na	na	na	na	na	na	na	na
GCCGTGAAAGAAAACAA CA (SEQ ID NO: 1175)	TATGAACCACTAAGGCGTcaacctgtgGCCGTGAAAAGAA AACAACAgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2748)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACTGTTGTTT TGTTTTCACGTGGCCAACGAAGGTCGGAGCTCCTTAGTTGGTTCATA (SEQ ID NO: 4321)	MRE11A_94226944.2 3-P1P2	MRE11A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAATTCCCGCGGGAACA GAA (SEQ ID NO: 1176)	TATGAACCACTAAGGCGTcaacctgtgATCGAATTCCGCGCGC GAAGAAGAgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2749)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACTTCGTCTCC CGCGGAATTCCACCAACGAAGGTCGGAGCTCCTTAGTTGGTTCATA (SEQ ID NO: 4322)	MRE11A_94227004.2 3-P1P2	MRE11A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGCTGGCGCGCTCAACC GAC (SEQ ID NO: 1177)	TATGAACCACTAAGGCGTcaacctgtgCACGCCTGGCGCGCT CACCGACgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2750)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACGTCGTGTC AGCGGCCAGTGCTCCAACAAGGTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4323)	MRGBP_+_6142616.23-P1P2	MRGBP	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGAAGCAAGGAGCCCC CGCC (SEQ ID NO: 1178)	TATGAACCACTAAGGCGTcaacctgtgCGAAGCAAGGAGCCCC GCTCGCTTGCTCGCCAACAAGGTCGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4324)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACGCGCGGG GCTCGCTTGCTCGCCAACAAGGTCGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4325)	MRGBP_+_6142799.23-P1P2	MRGBP	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTTGTGGCCCAGCATG GCG (SEQ ID NO: 1179)	TATGAACCACTAAGGCGTcaacctgtgGCTTGTGGCCCAG CATGGCGgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2752)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACCGCCATG CTGGGCCACAACGCCAACAAGGTCGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4326)	MRGBP_+_6142794.23-P1P2	MRGBP	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGTGAGGAGGTTTCG ACA (SEQ ID NO: 1180)	TATGAAGCACATAAGGCGTcaacctgtgGAGGTGAGGAGGTT TTCGACAgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2753)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACTGTCGAAA CGCTCCTCACCTCCAACAAGGTCGGAGCTCCTTAGTTGGTTCATA (SEQ ID NO: 4327)	MSH2_-47633633.2 3-P1P2	MSH2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGGTGTTCAAGACC CAG (SEQ ID NO: 1181)	TATGAACCACTAAGGCGTcaacctgtgGGGAAGGTGTTCAAA GAACCAGgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2754)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACCTGGGTTC TTCGAACACCTCCCCAACAAGGTCGGAGCGCTTAGTTGGTTCATA (SEQ ID NO: 4328)	MSH2_-47633615.2 3-P1P2	MSH2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTTCTTCAACCAAGAAGG TG (SEQ ID NO: 1182)	TATGAACCACTAAGGCGTcaacctgtgGTTTCTTCAACCAAG GAAGGTGgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2755)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACCACCTCGT GGTTGAAGAAACCACAACGAAGGTCGGAGCGCTTAGTTGGTTCATA (SEQ ID NO: 4329)	MSH2_-47633918.2 3-P1P2	MSH2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGAGACGCCTGGGAA CTG (SEQ ID NO: 1183)	TATGAAGCACATAAGGCGTcaacctgtgCGCGGAGACGCCTG GGTTGAAgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2756)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACCAGTTCCC AGGCGCTCTGCGCCAACAAGGTCGGAGCGCTTAGTTGGTTCATA (SEQ ID NO: 4330)	MSH3_-79953498.2 3-P1P2	MSH3	P1P2	DHFR_-79950488.2 3-P1P2	DHFR	P1P2	na	na	na	na	na	na	na	na	na	na
GGTTCCTGGCGAGACATTG GCA (SEQ ID NO: 1184)	TATGAACCACTAAGGCGTcaacctgtgGGTTCCTGGCGAGAA CATTGGCAgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2757)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACTGCCATGT CTTCGGCCGAAGCCAACAACGGTCGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4330)	MSH3_+_79950539.23-P1P2	MSH3	P1P2	DHFR_+_79950539.23-P1P2	DHFR	P1P2	na	na	na	na	na	na	na	na	na	na
GGCAGGCTTCCGGCCGAG ACA (SEQ ID NO: 1185)	TATGAACCACTAAGGCGTcaacctgtgGGCAGGCTTCCGG CGAGACAgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2758)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACTGTCTCGG CGGCAAGCCTCGCCCAACAAGGTCGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4331)	MSH3_+_79950544.23-P1P2	MSH3	P1P2	DHFR_+_79950544.23-P1P2	DHFR	P1P2	na	na	na	na	na	na	na	na	na	na
GCCCGTCGACCCGAGAAGT ATG (SEQ ID NO: 1186)	TATGAACCACTAAGGCGTcaacctgtgCGCCCTGACCCGAG AAGTATTgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2759)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACCATACTTC TCGGGTCGAGGGCCAACAAGGTCGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4332)	MSH4_+_76252635.23-P1P2	MSH4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGTCGCAGCTTCGTGTA GTT (SEQ ID NO: 1187)	TATGAACCACTAAGGCGTcaacctgtgGCGTCGGGCAGGTTC TGTAGTTgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2760)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACAAGTACAG AAAGCTGCGCGCCAACAAGGTCGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4333)	MSH4_-76262811.2 3-P1P2	MSH4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGTAGTTGGCGCTACTTC GAG (SEQ ID NO: 1188)	TATGAACCACTAAGGCGTcaacctgtgATGTAGTTGGCGCTA CTGGAGgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2761)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACCTTCGAGTA GCCGAACTACACGAACAAGGTCGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4334)	MSH4_-76262633.2 3-P1P2	MSH4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACCTGAGTTCGCTACA TGT (SEQ ID NO: 1189)	TATGAAGCACATAAGGCGTcaacctgtgGAACCTGAGTTCGC TACATGTgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2762)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACAGCCTGTAG CGAACTCAGGTTCCAACAAGGTCGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4335)	MSH5_+_31707820.23-P1P2	MSH5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACAGACGAGGTCCGGCG GTA (SEQ ID NO: 1190)	TATGAACCACTAAGGCGTcaacctgtgCACAGACGAGGTCCG GCGgttaagtgcctaagtgcctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 2763)	TCACCAGCACAGACTTAAGGCGTTAGCTAGCTTAAACTACCCGCG GCCTCCTCTGTGCCAACGAGGTCGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4336)	MSH5_+_31708167.23-P1P2	MSH5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na

GCAGTAACCTGAGTCGCTAC (SEQ ID NO: 1191)	TATGAACCACTAAGGCGTccaccttgttgGCAGTAACCTGAGTCGCTACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2764)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTAGCGACTCAGGTTACTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4337)	1	MSH5_+_31707824.23-P1P2	MSH5	P1P2	na	na	na	na	na	na	na	na	na
GAGGAGCCGCGCGGTAGATG (SEQ ID NO: 1192)	TATGAACCACTAAGGCGTccaccttgttgGAGGAGCCGCGCGGTAGATGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2765)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCATCTACCGCGCGGCTCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4338)	1	MSH6_-_48010319.23-P1P2	MSH6	P1P2	na	na	na	na	na	na	na	na	na
GGGAGCTCCGTCCGACAGAA (SEQ ID NO: 1193)	TATGAACCACTAAGGCGTccaccttgttgGGGAGCTCCGTCCGACAGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2766)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCTGTCGGACGGAGCTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4339)	1	MSH6_-_48010349.23-P1P2	MSH6	P1P2	na	na	na	na	na	na	na	na	na
GTCCGTCCGACAGAACGGTT (SEQ ID NO: 1194)	TATGAACCACTAAGGCGTccaccttgttgGTCCGTCCGACAGAACGGTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2767)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAACCGTTCTGTCGGACGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4340)	1	MSH6_-_48010354.23-P1P2	MSH6	P1P2	na	na	na	na	na	na	na	na	na
GAAGGAGCCGGGGGCCGCGA (SEQ ID NO: 1195)	TATGAACCACTAAGGCGTccaccttgttgGaaggagccgggggccgcgagtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2768)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCGGCCCCCGGCTCCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4341)	1	MTA1_+_105886510.23-P1P2	MTA1	P1P2	na	na	na	na	na	na	na	na	na
GGCCGGGCATAAAGGCGGTG (SEQ ID NO: 1196)	TATGAACCACTAAGGCGTccaccttgttgGgccgggcATAAAGGCGGTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2769)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCACCGCCTTTATGCCCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4342)	1	MTA1_+_105886406.23-P1P2	MTA1	P1P2	na	na	na	na	na	na	na	na	na
GTCGGGGCCGGGCATAAAGG (SEQ ID NO: 1197)	TATGAACCACTAAGGCGTccaccttgttgGtcggggccgggcATAAAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2770)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTTTATGCCCGGCCCCGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4343)	1	MTA1_+_105886411.23-P1P2	MTA1	P1P2	na	na	na	na	na	na	na	na	na
GCGGAAGGAGGCGAACCCTG (SEQ ID NO: 1198)	TATGAACCACTAAGGCGTccaccttgttgGCGGAAGGAGGCGAACCCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2771)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGGGTTCGCCTCCTTCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4344)	1	MTA2_+_62369279.23-P1P2	MTA2	P1P2	na	na	na	na	na	na	na	na	na
GGGATGAACCGCGACGGCTG (SEQ ID NO: 1199)	TATGAACCACTAAGGCGTccaccttgttgGGGATGAACCGCGACGGCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2772)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGCCGTCGCGGTTCATCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4345)	1	MTA2_+_62369066.23-P1P2	MTA2	P1P2	na	na	na	na	na	na	na	na	na
GGGGCCGGGATGAACCGCGA (SEQ ID NO: 1200)	TATGAACCACTAAGGCGTccaccttgttgGGGGCCGGGATGAACCGCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2773)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCGGTTCATCCCGGCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4346)	1	MTA2_+_62369072.23-P1P2	MTA2	P1P2	na	na	na	na	na	na	na	na	na
GCATGTACCGGGTCGGAGGT (SEQ ID NO: 1201)	TATGAACCACTAAGGCGTccaccttgttgGCATGTACCGGGTCGGAGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2774)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCTCCGACCCGGTACATGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4347)	1	MTA3_-_42795917.23-P1P2	MTA3	P1P2	na	na	na	na	na	na	na	na	na
GCTGCTGCCTCGCCACGGGG (SEQ ID NO: 1202)	TATGAACCACTAAGGCGTccaccttgttgGctgctgccTCGCCACggggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2775)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCCGTGGCGAGGCAGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4348)	1	MTA3_+_42795764.23-P1P2	MTA3	P1P2	na	na	na	na	na	na	na	na	na
GGCCAACTGCCGGGAGCCAA (SEQ ID NO: 1203)	TATGAACCACTAAGGCGTccaccttgttgGGCCAACTGCCGGGAGCCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2776)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGGCTCCCGGCAGTTGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4349)	1	MTA3_-_42796005.23-P1P2	MTA3	P1P2	na	na	na	na	na	na	na	na	na
GACGGGGCCCCAAACACAGG (SEQ ID NO: 1204)	TATGAACCACTAAGGCGTccaccttgttgGACGGGGCCCCAAACACAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2777)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTGTGTTTGGGGCCCCGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4350)	1	MUS81_+_65628068.23-P1P2	MUS81	P1P2	na	na	na	na	na	na	na	na	na
GCACAGGGGGCACTAACGAG (SEQ ID NO: 1205)	TATGAACCACTAAGGCGTccaccttgttgGCACAGGGGGCACTAACGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2778)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCGTTAGTGCCCCCTGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4351)	1	MUS81_+_65628055.23-P1P2	MUS81	P1P2	na	na	na	na	na	na	na	na	na
GTGGGGCCCCGTGATCTCAA (SEQ ID NO: 1206)	TATGAACCACTAAGGCGTccaccttgttgGTGGGGCCCCGTGATCTCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2779)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGAGATCACGGGGCCCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4352)	1	MUS81_-_65628100.23-P1P2	MUS81	P1P2	na	na	na	na	na	na	na	na	na
GTGAAGACAGCAGAACACGG (SEQ ID NO: 1207)	TATGAACCACTAAGGCGTccaccttgttgGTGAAGACAGCAGAACACGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2780)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGTGTTCTGCTGTCTTCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4353)	1	MUTYH_-_45805607.23-P1P2	MUTYH	P1P2	TOE1_-_45805607.23-P1P2	TOE1	P1P2	na	na	na	na	na	na

GCTCGGGCTCATAGTTCT AG (SEQ ID NO: 1208)	TATGAAGCACTAAGGCGTcaacttgtgGCTCGGGCTCATA ATGAGCGCCGAGCCAACAAG3'OTGGACGGCTTAGTGCTGATA (SEQ ID NO: 4354)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCTAGAACT		MUTYH_-_45880855.2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGCTCCTCAAGCTT CAG (SEQ ID NO: 1209)	TATGAACCACTAAGGCGTcaacttgtgGAGAGCTCCTCAA TGAGGAAGCGTCCGCAAGTCTGCTGGTGA (SEQ ID NO: 4355)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCTGAAGCT		MUTYH_-_45880873.2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GCAAAAGCTTCCCACCCTC CC (SEQ ID NO: 1210)	TATGAACCACTAAGGCGTcaacttgtgCAAAAGCTTCCCAC TGGGAAGCTTTGCCAACAAGGTGGACGCTTAGTGGTTCATA (SEQ ID NO: 4356)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGAGGG		NABP1_+_19: 2543255.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GGCACAGCGGGCCGGTG GAGT (SEQ ID NO: 1211)	TATGAACCACTAAGGCGTcaacttgtgCGCACAGCGGGCGC GTCGGAGTgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4357)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACACTCCACG		NABP1_+_19: 2543448.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GTGTCACAGCCAAACGG GAG (SEQ ID NO: 1212)	TATGAACCACTAAGGCGTcaacttgtgTGTGACAGCCAAA ACGGGAAgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4358)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCGTT		NABP1_+_19: 2542890.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GAGCCGAGTGTAGACAGG GGT (SEQ ID NO: 1213)	TATGAACCACTAAGGCGTcaacttgtgGAGCCGAGTGTAGA CAAGGGGTgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4359)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACACCCCTGT		NABP2_+_56: 6183839.23-P1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GCGTGCGCGGTGGCTTG GGT (SEQ ID NO: 1214)	TATGAAGCACTAAGGCGTcaacttgtgCGCTGCGCGGCTGG CTTGGCTgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4360)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACACCCAAG		NABP2_-_58618228.2 3-P1	NABP2	na	na	na	na	na	na	na	na	na	na	na	na
GTCCACGGGGGCACATC CGG (SEQ ID NO: 1215)	TATGAACCACTAAGGCGTcaacttgtgGTCCACGGGGGCAG CATCCCGgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4361)		TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACCCCGATG		NABP2_-_58618158.2 3-P1	P1	na	na	na	na	na	na	na	na	na	na	na	na
GAGCCTTCGTCCCTCCTC GT (SEQ ID NO: 1216)	TATGAACCACTAAGGCGTcaacttgtgGAGCCTTCGTCCCT CCTCGTgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4362)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACACGAGGA		NABP2_+_56: 6159810.23-P2	P2	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGAGCTGCCTACGGA GGA (SEQ ID NO: 1217)	TATGAAGCACTAAGGCGTcaacttgtgCGCAGGAGCTGCCT ACGAGGAgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4363)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTCCTCGTA		NABP2_-_58615822.2 3-P2	NABP2	na	na	na	na	na	na	na	na	na	na	na	na
GTGCCTACGAAGAGGGGA CGA (SEQ ID NO: 1218)	TATGAACCACTAAGGCGTcaacttgtgGTGCCTACGAAGGA CGGGAGCgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4364)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTGTCCCT		NABP2_-_58615839.2 3-P2	NABP2	na	na	na	na	na	na	na	na	na	na	na	na
GCCAGCGCTTACTACGTGA GC (SEQ ID NO: 1219)	TATGAACCACTAAGGCGTcaacttgtgGCCAGCGCTTACTAC GTAGTAAGCCCTCGGCCAACAAGGTGGACGCGCTTAGTGGTTCATA (SEQ ID NO: 4365)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGGGTCACG		NAT10_-_34127225.2 3-P1P2	NAT10	na	na	na	na	na	na	na	na	na	na	na	na
GGAACCACGAGGCATACGC TAG (SEQ ID NO: 1220)	TATGAACCACTAAGGCGTcaacttgtgCGACCACGAGGCATA TGCCTAGgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4366)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCTAGCGTA		NAT10_-_34127248.2 3-P1P2	NAT10	na	na	na	na	na	na	na	na	na	na	na	na
GTGTCTACCAGTTCCTGA CA (SEQ ID NO: 1221)	TATGAACCACTAAGGCGTcaacttgtgGTGTCTACCAGTTC CTGACAGgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4367)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTGTCAGGA		NAT10_-_34127183.2 3-P1P2	NAT10	na	na	na	na	na	na	na	na	na	na	na	na
GACATCGGTTCCGGGCTCC TCA (SEQ ID NO: 1222)	TATGAACCACTAAGGCGTcaacttgtgGACATCGGTTCCGG GCTCCTCAgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4368)		TCACCACGCAGACTTTACGCAGCTTAGCTCTTAAACTGAGGAG		NBN_-_90956805.2 3-P1P2	NBN	P1P2	na	na	na	na	na	na	na	na	na	na	na
GACGTCGCAACCGGTGTAA CCG (SEQ ID NO: 1223)	TATGAACCACTAAGGCGTcaacttgtgGACGTCGCAACCG GTGTAACCGgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4369)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGTTAC		NBN_-_90956837.2 3-P1P2	NBN	P1P2	na	na	na	na	na	na	na	na	na	na	na
GTGCTAGAGCGAGCGCGG ATA (SEQ ID NO: 1224)	TATGAACCACTAAGGCGTcaacttgtgGTGCTAGAGCGAG CGCGGATAgttaagngcaaegtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 4370)		TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTATCGCGC		NBN_-_90956863.2 3-P1P2	NBN	P1P2	na	na	na	na	na	na	na	na	na	na	na

GAACAAGTAACCCAGGCGCC (SEQ ID NO: 1225)	TATGAACCACTAAGGCGTccaccttgttgGAACAAGTAACCCAGGCGCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2798)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGCGCCTGGGTTACTTGTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4371)	1	non-targeting_00249	negative_control	na	na	na	na	na	na	na	na	na	na
GACCGCGGGCGTAAGCCCCC (SEQ ID NO: 1226)	TATGAACCACTAAGGCGTccaccttgttgGACCGCGGGCGTAAGCCCCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2799)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGGGCTTACGCCCGCGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4372)	1	non-targeting_00364	negative_control	na	na	na	na	na	na	na	na	na	na
GACCGCTGCACGGGTTGGCG (SEQ ID NO: 1227)	TATGAACCACTAAGGCGTccaccttgttgGACCGCTGCACGGGTTGGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2800)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCAACCCGTGCAGCGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4373)	1	non-targeting_03019	negative_control	na	na	na	na	na	na	na	na	na	na
GACTCATAGTCGAACAGTGC (SEQ ID NO: 1228)	TATGAACCACTAAGGCGTccaccttgttgGACTCATAGTCGAACAGTGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2801)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCACTGTTCGACTATGAGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4374)	1	non-targeting_03292	negative_control	na	na	na	na	na	na	na	na	na	na
GAGATCAATATTATGCAAGA (SEQ ID NO: 1229)	TATGAACCACTAAGGCGTccaccttgttgGAGATCAATATTATGCAAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2802)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTTGCATAATATTGATCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4375)	1	non-targeting_01358	negative_control	na	na	na	na	na	na	na	na	na	na
GAGCCTCTTGGTTCCATGGA (SEQ ID NO: 1230)	TATGAACCACTAAGGCGTccaccttgttgGAGCCTCTTGGTTCCATGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2803)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCATGGAACCAAGAGGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4376)	1	non-targeting_00026	negative_control	na	na	na	na	na	na	na	na	na	na
GAGGCGCCTTTAATTTGCGT (SEQ ID NO: 1231)	TATGAACCACTAAGGCGTccaccttgttgGAGGCGCCTTTAATTTGCGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2804)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGCAAATTAAAGGCGCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4377)	1	non-targeting_03346	negative_control	na	na	na	na	na	na	na	na	na	na
GAGGGTAACGCAGAAGAAGG (SEQ ID NO: 1232)	TATGAACCACTAAGGCGTccaccttgttgGAGGGTAACGCAGAAGAAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2805)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTTCTTCTGCGTTACCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4378)	1	non-targeting_02964	negative_control	na	na	na	na	na	na	na	na	na	na
GAGGTGCCCCGTCGCCGGGG (SEQ ID NO: 1233)	TATGAACCACTAAGGCGTccaccttgttgGAGGTGCCCCGTCGCCGGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2806)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCCGGCGACGGGGCACCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4379)	1	non-targeting_01091	negative_control	na	na	na	na	na	na	na	na	na	na
GAGTCGCGCAACGCGCGGCC (SEQ ID NO: 1234)	TATGAACCACTAAGGCGTccaccttgttgGAGTCGCGCAACGCGCGGCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2807)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGCCGCGCGTTGCGCGACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4380)	1	non-targeting_02692	negative_control	na	na	na	na	na	na	na	na	na	na
GCAGCCCCCAGAGAAAGAGC (SEQ ID NO: 1235)	TATGAACCACTAAGGCGTccaccttgttgGCAGCCCCCAGAGAAAGAGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2808)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCTCTTTCTCTGGGGGCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4381)	1	non-targeting_01631	negative_control	na	na	na	na	na	na	na	na	na	na
GCAGGCCGCCCTGCTAAGAA (SEQ ID NO: 1236)	TATGAACCACTAAGGCGTccaccttgttgGCAGGCCGCCCTGCTAAGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2809)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCTTAGCAGGGCGGCCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4382)	1	non-targeting_01439	negative_control	na	na	na	na	na	na	na	na	na	na
GCAGTCCTCGGCGGAGGCGT (SEQ ID NO: 1237)	TATGAACCACTAAGGCGTccaccttgttgGCAGTCCTCGGCGGAGGCGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2810)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGCCTCCGCCGAGGACTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4383)	1	non-targeting_03051	negative_control	na	na	na	na	na	na	na	na	na	na
GCCAACGCTAGATATAAACA (SEQ ID NO: 1238)	TATGAACCACTAAGGCGTccaccttgttgGCCAACGCTAGATATAAACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2811)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTTTATATCTAGCGTTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4384)	1	non-targeting_01959	negative_control	na	na	na	na	na	na	na	na	na	na
GCCCCGCGGAAAGGCGTCTC (SEQ ID NO: 1239)	TATGAACCACTAAGGCGTccaccttgttgGCCCCGCGGAAAGGCGTCTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2812)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGACGCCTTTCCGCGGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4385)	1	non-targeting_03059	negative_control	na	na	na	na	na	na	na	na	na	na
GCCCCTCGATTCCCGAACAG (SEQ ID NO: 1240)	TATGAACCACTAAGGCGTccaccttgttgGCCCCTCGATTCCCGAACAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2813)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGTTCGGGAATCGAGGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4386)	1	non-targeting_00398	negative_control	na	na	na	na	na	na	na	na	na	na
GCCGAAGCTCAAGCGTAAGA (SEQ ID NO: 1241)	TATGAACCACTAAGGCGTccaccttgttgGCCGAAGCTCAAGCGTAAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2814)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTTACGCTTGAGCTTCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4387)	1	non-targeting_01332	negative_control	na	na	na	na	na	na	na	na	na	na

GCCGGCAGTTGGGCAGTTTG (SEQ ID NO: 1242)	TATGAACCACTAAGGCGTccaccttgttgGCCGGCAGTTGGGCAGTTTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2815)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAAACTGCCCAACTGCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4388)	1	non-targeting_00566	negative_control	na	na	na	na	na	na	na	na	na	na
GCCTTACGTGGGATTCGGTT (SEQ ID NO: 1243)	TATGAACCACTAAGGCGTccaccttgttgGCCTTACGTGGGATTCGGTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2816)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAACCGAATCCCACGTAAGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4389)	1	non-targeting_02359	negative_control	na	na	na	na	na	na	na	na	na	na
GCGACATCGGTTACATGTGG (SEQ ID NO: 1244)	TATGAACCACTAAGGCGTccaccttgttgGCGACATCGGTTACATGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2817)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACATGTAACCGATGTCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4390)	1	non-targeting_00097	negative_control	na	na	na	na	na	na	na	na	na	na
GCGCCGGAATAGCACATCAA (SEQ ID NO: 1245)	TATGAACCACTAAGGCGTccaccttgttgGCGCCGGAATAGCACATCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2818)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGATGTGCTATTCCGGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4391)	1	non-targeting_00678	negative_control	na	na	na	na	na	na	na	na	na	na
GCGCGAAGACGCGACACCAG (SEQ ID NO: 1246)	TATGAACCACTAAGGCGTccaccttgttgGCGCGAAGACGCGACACCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2819)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGGTGTCGCGTCTTCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4392)	1	non-targeting_03086	negative_control	na	na	na	na	na	na	na	na	na	na
GCGCGCCCATTCGGACTGTC (SEQ ID NO: 1247)	TATGAACCACTAAGGCGTccaccttgttgGCGCGCCCATTCGGACTGTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2820)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGACAGTCCGAATGGGCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4393)	1	non-targeting_02293	negative_control	na	na	na	na	na	na	na	na	na	na
GCGGATTCCCCCACGGCGA (SEQ ID NO: 1248)	TATGAACCACTAAGGCGTccaccttgttgGCGGATTCCCCCACGGCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2821)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCCGTGGGGGGAATCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4394)	1	non-targeting_01222	negative_control	na	na	na	na	na	na	na	na	na	na
GCGGGGATGTCCACACTCGG (SEQ ID NO: 1249)	TATGAACCACTAAGGCGTccaccttgttgGCGGGGATGTCCACACTCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2822)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGAGTGTGGACATCCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4395)	1	non-targeting_01566	negative_control	na	na	na	na	na	na	na	na	na	na
GCGTCGGGGAAGCCAACCTA (SEQ ID NO: 1250)	TATGAACCACTAAGGCGTccaccttgttgGCGTCGGGGAAGCCAACCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2823)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGGTTGGCTTCCCCGACGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4396)	1	non-targeting_03264	negative_control	na	na	na	na	na	na	na	na	na	na
GCTGAGATGACTCAACCACG (SEQ ID NO: 1251)	TATGAACCACTAAGGCGTccaccttgttgGCTGAGATGACTCAACCACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2824)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTGGTTGAGTCATCTCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4397)	1	non-targeting_02937	negative_control	na	na	na	na	na	na	na	na	na	na
GCTGATACCGTAAATGACCC (SEQ ID NO: 1252)	TATGAACCACTAAGGCGTccaccttgttgGCTGATACCGTAAATGACCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2825)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGTCATTTACGGTATCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4398)	1	non-targeting_01042	negative_control	na	na	na	na	na	na	na	na	na	na
GCTGCACTACTAAACCTCCG (SEQ ID NO: 1253)	TATGAACCACTAAGGCGTccaccttgttgGCTGCACTACTAAACCTCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2826)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGAGGTTTAGTAGTGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4399)	1	non-targeting_02821	negative_control	na	na	na	na	na	na	na	na	na	na
GCTTTATTACCGCAAGGAGA (SEQ ID NO: 1254)	TATGAACCACTAAGGCGTccaccttgttgGCTTTATTACCGCAAGGAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2827)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTCCTTGCGGTAATAAAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4400)	1	non-targeting_01528	negative_control	na	na	na	na	na	na	na	na	na	na
GGAAACGACAGGGAACCCGT (SEQ ID NO: 1255)	TATGAACCACTAAGGCGTccaccttgttgGGAAACGACAGGGAACCCGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2828)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGGGTTCCCTGTCGTTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4401)	1	non-targeting_00869	negative_control	na	na	na	na	na	na	na	na	na	na
GGACGGCCCAGAGTACAATA (SEQ ID NO: 1256)	TATGAACCACTAAGGCGTccaccttgttgGGACGGCCCAGAGTACAATAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2829)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTATTGTACTCTGGGCCGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4402)	1	non-targeting_00515	negative_control	na	na	na	na	na	na	na	na	na	na
GGACTACTCGTGGCGCTTGC (SEQ ID NO: 1257)	TATGAACCACTAAGGCGTccaccttgttgGGACTACTCGTGGCGCTTGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2830)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCAAGCGCCACGAGTAGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4403)	1	non-targeting_00155	negative_control	na	na	na	na	na	na	na	na	na	na
GGATCGCGCCACGAAAATAG (SEQ ID NO: 1258)	TATGAACCACTAAGGCGTccaccttgttgGGATCGCGCCACGAAAATAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2831)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTATTTTCGTGGCGCGATCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4404)	1	non-targeting_02607	negative_control	na	na	na	na	na	na	na	na	na	na

GGCAGCGACCGGCATGCCGC (SEQ ID NO: 1259)	TATGAACCACTAAGGCGTccaccttgttgGGCAGCGACCGGCATGCCGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2832)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGGCATGCCGGTCGCTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4405)	1	non-targeting_00240	negative_control	na	na	na	na	na	na	na	na	na	na
GGCCGCCAAACCGATCACGT (SEQ ID NO: 1260)	TATGAACCACTAAGGCGTccaccttgttgGGCCGCCAAACCGATCACGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2833)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGTGATCGGTTTGGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4406)	1	non-targeting_03451	negative_control	na	na	na	na	na	na	na	na	na	na
GGCGTCGGCGCGACACTGCC (SEQ ID NO: 1261)	TATGAACCACTAAGGCGTccaccttgttgGGCGTCGGCGCGACACTGCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2834)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGCAGTGTCGCGCCGACGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4407)	1	non-targeting_02822	negative_control	na	na	na	na	na	na	na	na	na	na
GGCTGCGTGCGCCGCGCACG (SEQ ID NO: 1262)	TATGAACCACTAAGGCGTccaccttgttgGGCTGCGTGCGCCGCGCACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2835)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTGCGCGGCGCACGCAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4408)	1	non-targeting_00745	negative_control	na	na	na	na	na	na	na	na	na	na
GGCTTGTCGAGTCCGAGCAA (SEQ ID NO: 1263)	TATGAACCACTAAGGCGTccaccttgttgGGCTTGTCGAGTCCGAGCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2836)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGCTCGGACTCGACAAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4409)	1	non-targeting_02305	negative_control	na	na	na	na	na	na	na	na	na	na
GGGCTCCAGTGTAAAACCCG (SEQ ID NO: 1264)	TATGAACCACTAAGGCGTccaccttgttgGGGCTCCAGTGTAAAACCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2837)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGTTTTACACTGGAGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4410)	1	non-targeting_00889	negative_control	na	na	na	na	na	na	na	na	na	na
GGGGGATGGGAGACAATTGG (SEQ ID NO: 1265)	TATGAACCACTAAGGCGTccaccttgttgGGGGGATGGGAGACAATTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2838)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCAATTGTCTCCCATCCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4411)	1	non-targeting_03512	negative_control	na	na	na	na	na	na	na	na	na	na
GGGGTCGAATCCACACGGTA (SEQ ID NO: 1266)	TATGAACCACTAAGGCGTccaccttgttgGGGGTCGAATCCACACGGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2839)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACCGTGTGGATTCGACCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4412)	1	non-targeting_00740	negative_control	na	na	na	na	na	na	na	na	na	na
GGGTAGCTTCGCTCCATCAG (SEQ ID NO: 1267)	TATGAACCACTAAGGCGTccaccttgttgGGGTAGCTTCGCTCCATCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2840)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGATGGAGCGAAGCTACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4413)	1	non-targeting_03260	negative_control	na	na	na	na	na	na	na	na	na	na
GGGTCGCACTGGGAACAACA (SEQ ID NO: 1268)	TATGAACCACTAAGGCGTccaccttgttgGGGTCGCACTGGGAACAACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2841)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTTGTTCCCAGTGCGACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4414)	1	non-targeting_02085	negative_control	na	na	na	na	na	na	na	na	na	na
GGGTCTCGTACCGGCTGGTC (SEQ ID NO: 1269)	TATGAACCACTAAGGCGTccaccttgttgGGGTCTCGTACCGGCTGGTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2842)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGACCAGCCGGTACGAGACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4415)	1	non-targeting_03178	negative_control	na	na	na	na	na	na	na	na	na	na
GGGTGGGTCACGCGGAAGGG (SEQ ID NO: 1270)	TATGAACCACTAAGGCGTccaccttgttgGGGTGGGTCACGCGGAAGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2843)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCTTCCGCGTGACCCACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4416)	1	non-targeting_00074	negative_control	na	na	na	na	na	na	na	na	na	na
GGTCCAGGCACACTACCGGC (SEQ ID NO: 1271)	TATGAACCACTAAGGCGTccaccttgttgGGTCCAGGCACACTACCGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2844)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCGGTAGTGTGCCTGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4417)	1	non-targeting_02613	negative_control	na	na	na	na	na	na	na	na	na	na
GGTCCGAACAGCCCAGGCTC (SEQ ID NO: 1272)	TATGAACCACTAAGGCGTccaccttgttgGGTCCGAACAGCCCAGGCTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2845)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGCCTGGGCTGTTCGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4418)	1	non-targeting_03460	negative_control	na	na	na	na	na	na	na	na	na	na
GGTCCGGAACCGCCCGAGGG (SEQ ID NO: 1273)	TATGAACCACTAAGGCGTccaccttgttgGGTCCGGAACCGCCCGAGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2846)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCTCGGGCGGTTCCGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4419)	1	non-targeting_03220	negative_control	na	na	na	na	na	na	na	na	na	na
GGTCTTTCGAAGGCAGCTTT (SEQ ID NO: 1274)	TATGAACCACTAAGGCGTccaccttgttgGGTCTTTCGAAGGCAGCTTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2847)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAAAGCTGCCTTCGAAAGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4420)	1	non-targeting_03612	negative_control	na	na	na	na	na	na	na	na	na	na
GGTTGGCTGTACTGGGCCAA (SEQ ID NO: 1275)	TATGAACCACTAAGGCGTccaccttgttgGGTTGGCTGTACTGGGCCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2848)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGGCCCAGTACAGCCAACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4421)	1	non-targeting_00948	negative_control	na	na	na	na	na	na	na	na	na	na

GTAACGACGACCCGTCA GGG (SEQ ID NO: 1275)	TATGAACACACTAAG3GCGTcaacctgtgG7AAGGACGACCC G3GGTCGTCGTTAACAACAAG3GTGGACCGCCTTAGTG3TTCATA (SEQ ID NO: 2849)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACCCGTGAC 13	non-targeting_005	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCACCAGCGCATAAGC GCA (SEQ ID NO: 1277)	TATGAACACACTAAGGCGTcaacctgtgG7CACCAGCGCATA T3CCGTCGTTG3AGCCAACAAG3GTTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 2850)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACTGCGCTTA 84	non-targeting_009	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCTTGGCCCGCAGATC GCT (SEQ ID NO: 1278)	TATGAACACACTAAGGCGTcaacctgtgG7CCTCGGCCCGCA GATCGCGTgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2851)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACCATCCATCGT 42	non-targeting_029	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGTCGAGACGAGACGCG GGC (SEQ ID NO: 1279)	TATGAACACACTAAGGCGTcaacctgtgG7TCTCGAGACGCA GACGCGGGgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2852)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACGCCGCTC 76	non-targeting_036	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGAGTCGCGCTCTGT CGG (SEQ ID NO: 1280)	TATGAACACACTAAGGCGTcaacctgtgG7GAAGTCCGCGCT CTGTCGGgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2853)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACCGAACAG AG3GTCGACTTCCACCAAGAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4426)	non-targeting_020	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGCCCCCCGATCGTA GGT (SEQ ID NO: 1281)	TATGAACACACTAAGGCGTcaacctgtgG7GGCCCCCCGGAT CGTAGGTgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2854)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACACCTACGA TCGGGGGGGCCACCAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4427)	non-targeting_030	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGGTTGAACAAGCCAAT AG (SEQ ID NO: 1282)	TATGAACACACTAAGGCGTcaacctgtgG7TGGTTCGAACAAG CCAATAGgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2855)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACTCTATTGCC TTGTTCAACCACCAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4428)	non-targeting_026	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTAGGGCGTTAAG3GTT CGA (SEQ ID NO: 1283)	TATGAACACACTAAGGCGTcaacctgtgG7TAGGGCG7TAAG GCGTTCGAgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2856)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACTCGAACCC TTATGCCCTAACCAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4429)	non-targeting_015	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTGGTTACCCCTCGATT GG (SEQ ID NO: 1284)	TATGAACACACTAAGGCGTcaacctgtgG7TGGTTACCCCTC GATTGGgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2857)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACCCAATCGA GGGGTAACCAACCACAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4430)	non-targeting_038	negative_control	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGCTAGCG3TGGCGACAG GGT (SEQ ID NO: 1285)	TATGAAGCACACTAAG3GCGTcaacctgtgGAGCTAGCG7GGC ACAGG3GTgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2858)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACCACCTGTC CCACGCTAGCTCCAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4431)	NEIL1_+_75639678.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAACAGAACCATCTCAA T3T (SEQ ID NO: 1286)	TATGAACACACTAAG3GCGTcaacctgtgGCAACAGAACCATC T3G3TTCTG7TTG3CAACAAG3GTCGGTTGA (SEQ ID NO: 2859)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACACTTGAGA NEIL1_-_75639455.23-P1P2 (SEQ ID NO: 4432)	NEIL1	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGGTACGCACAGGCC CTA (SEQ ID NO: 1287)	TATGAACACACTAAGGCGTcaacctgtgG3CGGTACGGCACA G3CCCCTAgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2860)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACTAGGGGCC TGTGCGTACCG3CCCAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4433)	NEIL1_+_75639581.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGCGCAGATCTACC GGA (SEQ ID NO: 1288)	TATGAAGCACACTAAG3GCGTcaacctgtgCAG3GCCGGCAGATC TACCGGAgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2861)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACTCCG3CTA GATCTGGG3GCCTCCAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4434)	NEIL2_+_116273317.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCACAGCCGCTCCACGC TCG (SEQ ID NO: 1289)	TATGAACACACTAAGGCGTcaacctgtgG3CACAGCCGCTTC CAG3CTG3gttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2862)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACGCCAGCGTG GAG3CCGCTGTG3CCCAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4435)	NEIL2_+_11627227.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCGGCGCTGCTGCACT CAC (SEQ ID NO: 1290)	TATGAACACACTAAGGCGTcaacctgtgG3GCGCGGCGCTGC CACTCAGgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2863)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACGTGAGTG CAGCGCGCGGCGCTGCTGCACTCAGTCGTGGTTCATA (SEQ ID NO: 4436)	NEIL2_+_11627297.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGACGGCCTGTGCAATC GGT (SEQ ID NO: 1291)	TATGAACACACTAAG3GCGTcaacctgtgGACGACGGCCTGC AATCGGTgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2864)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACACCG3ATTG CAG3GCGCTG3TCCAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4437)	NEIL3_+_178231698.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CATGTCTGTGGCCGCGCA CTG (SEQ ID NO: 1292)	TATGAACACACTAAGGCGTcaacctgtgCATGTCTGTG3CCGCA CTGCGCACTgttaagagtcaagagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2865)	TCACCACGCAGACTTTACGCAGCTTAAGCTTGTTAAACGCAGTGCC GGCCACACGAGATGCCACAACAAG3GTGGACCGCTTAGTG3TTCATA (SEQ ID NO: 4438)	NEIL3_+_178231057.23-P1P2	PP1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTCTCACGACGGCCGTGC AAT (SEQ ID NO: 1293)	TATGAAGCACTAAGGCGTcaacttgtggGTCTCACGACGGCC GTGCGTAAgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2656)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAAGATTGCAG 39CGTGCTGGTGAGACCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4438)	NEIL3_-_17823182 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGAAGGGCTGCGGGTT AGG (SEQ ID NO: 1294)	TATGAACCACTAAGGCGTcaacttgtgGAGAAGGGCTGCG GGTTAGGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2657)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACCGTAACCC GCAGCCCTCTCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4440)	NFRKB_+_1_25766408.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGTGCGTCTCGGAGAC CCG (SEQ ID NO: 1295)	TATGAACCACTAAGGCGTcaacttgtgGAGTGCGTCTCGG AGACCCGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2658)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACCGGGAGTC GCAGAGGACGTCGCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4441)	NFRKB_+_1_25766434.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGGGGAGCGAGTGCG TCT (SEQ ID NO: 1296)	TATGAACCACTAAGGCGTcaacttgtgGTGGGGGAGCGAG TGCGTCTgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2659)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACAGACGCA CTCGGTCTCCGGCACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4442)	NFRKB_+_1_25766443.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGCTCGCGCAAACCC GAA (SEQ ID NO: 1297)	TATGAACCACTAAGGCGTcaacttgtgGCCCGCTCGCGCA AACCCGAAgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2570)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACTTCGGTTT GCGCGAGCGCGCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4443)	NHEJ1_-_22002564 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGAAAGCGCCTAGAGT AAG (SEQ ID NO: 1298)	TATGAACCACTAAGGCGTcaacttgtgGCCGAAAGCGCCTA GAGTAAGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2671)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACCTTACTCT AGGCGCTTTCGGCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4444)	NHEJ1_-_22002582823-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCGTGCGGCTAAGAG AGT (SEQ ID NO: 1299)	TATGAACCACTAAGGCGTcaacttgtgGTGCGTGCGGCTA AGAGAAGTgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2672)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACGCTCTCTT AGCGCGCACGCACCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4445)	NHEJ1_+_22200258723-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACCACCAAGCCGCGAG 93A (SEQ ID NO: 1300)	TATGAACCACTAAGGCGTcaacttgtgGCACCACCAAGCC GCGAGGCGAgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2673)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACTCCCTGC GGCTTGGTGCTCCACCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4446)	NIPBL_-_36937698223-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTTGAGTTCGGTTACCGAC GGA (SEQ ID NO: 1301)	TATGAACCACTAAGGCGTcaacttgtgGGTTGAGTTCGGT TACCGGAgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2674)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACTCCGTTCG GTACCGACTCACCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4447)	NIPBL_+_36879945.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTCGGTTACCGACGAGG GAG (SEQ ID NO: 1302)	TATGAAGCACTAAGGCGTcaacttgtgGTTCGGTTACCGA CGAGGGAGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2675)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACCTCCCTCG GTTCGGTACCGACGCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4448)	NIPBL_+_36879940.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGTTCATGCCGGAGCTC CTG (SEQ ID NO: 1303)	TATGAACCACTAAGGCGTcaacttgtgGCGGTTCATGCCGG AGCTCCTGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2676)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTTAAACGAGGAGT CCGGGCATCAACCGGCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4449)	NTHL1_-_20975383.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTCGTCCCTACACCCC CCG (SEQ ID NO: 1304)	TATGAACCACTAAGGCGTcaacttgtgGGCTCGTCCCTACA CCCCCCGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2677)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACCGGGGGGT GTAGGGAGGAGCCCAACAAGGTGGAGCGCCTTAGTGGTTCAT A (SEQ ID NO: 4450)	NTHL1_-_20977633.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCCTCTCCGGGAGAA GAG (SEQ ID NO: 1305)	TATGAACCACTAAGGCGTcaacttgtgGGGCCCTCTCCGG GAGAAGAGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2678)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACCTCTTCTC CGGGAGAGGGCCCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4451)	NTHL1_+_20977619.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAGCGCGCCCCGGG CCG (SEQ ID NO: 1306)	TATGAACCACTAAGGCGTcaacttgtgGGAGCGCGCCCCG GGCCGTGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2679)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACGCTTCCGG GGGCGTGGGCTCCCAACAAGGTGGAGCGCCTTAGTGGTTCAT A (SEQ ID NO: 4452)	NUDT1_-_23911882.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGTCGTGGTGACCAG GGA (SEQ ID NO: 1307)	TATGAACCACTAAGGCGTcaacttgtgGGAGTCGTGGTGA CCAGGGGAgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2680)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACTCCGTGGT CACCAGCGACTCCCAAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4453)	NUDT1_-_22911940.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCACGCCCCCGGAAG CGG (SEQ ID NO: 1308)	TATGAACCACTAAGGCGTcaacttgtgGTGCACGCCCCGG GAAGCGGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2681)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACCCGCTTCC GTGCGTGGGTGGGCCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4454)	NUDT1_-_22911835.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTAGGGTGCGCTAGT ACG (SEQ ID NO: 1309)	TATGAACCACTAAGGCGTcaacttgtgGGGTAGGGTGCGC TAGTACGgttaagtgctaagtgtgGTAAAGTCTGCTGGTGTA (SEQ ID NO: 2682)	TCACGACGCAGAGTTTAGCGAGGTTAGGTCTTAAACGCGTAGTA GCGCCGGCGTACCCCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4455)	OGG1_-_9791952.23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTACGATCGCCCATGCG GCT (SEQ ID NO: 1319)	TATGAAGCACTAAG3GCTcaacttgtgG7ACGATGCCCCAT (GCGCGTgttaagagcaagcgtGTAAA3TCTCGTGGTGA (SEQ ID NO: 4456)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACAGGCGCA		OGG1_+_97919911 23-P1P2	OGG1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCAGAGCGAGGCGCAGG GGA (SEQ ID NO: 1311)	TATGAACCACTAAGGCTcaacttgtgG7CAGAGCGAGGG (CGCTCGTCTGTGACCAACAAGGTTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4457)	TCACGACGCAGAAGTTTACGCAGCTTAGCTCTTAAACATCCCGTG		OGG1_+_97820433 23-P1P2	OGG1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
CAAGGCCGCCCACAGGAAG AAT (SEQ ID NO: 1312)	TATGAACCACTAAGGCGTcaacttgtgAGCGCGCCCCACGG (CTGGGTGCGCTCCAACAAGGTTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4458)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACATTCCCTTG		PALB2_-_23652662 23-P1P2	PALB2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GATTTAATTGGCCGGAGT TT (SEQ ID NO: 1313)	TATGAAGCACTAAGGGCGTcaacttgtgGATTTAATTGGCCGG	TCACGACGCAGAGTTTACGCAGCTTAGCTCTTAAACAAACTCCGG		PALB2_+_23652846 23-P1P2	PALB2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCTGAGCCGCGTTTCGC CA (SEQ ID NO: 1314)	TATGAACCACTAAGGCGTcaacttgtgGCTCTGAGCCGCGC	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTGGGAAA		PALB2_+_23652376 23-P1P2	PALB2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GACCATGGCTTGCCTGTTT AT (SEQ ID NO: 1315)	TATGAAGCACTAAGGGCGTcaacttgtgGACCATGCTTGCCT	TCACGACGCAGAGCTTTACGCAGCTTAGCTCTTAAACATAAACAG		PARG_+_51130346 23-ENST00000202038.3	PARG	ENST00000202038.3	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGACACAGGGCGCCATA AAC (SEQ ID NO: 1316)	TATGAAGCACTAAGGGCGTcaacttgtgGAGACACAGGGCC	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGTTTATGG		PARG_-_51130367 23-ENST00000202038.3	PARG	ENST00000202038.3	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTAGTAGAATGTCTCT AA (SEQ ID NO: 1317)	TATGAACCACTAAGGGCGTcaacttgtgGCCTAGTAGAATG	TCACCAGCAGAGACTTTACGCAGCTTAGCTCTTAAACTTAGAGAC		PARG_-_51130247 23-ENST00000202038.3	PARG	ENST00000202038.3	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTGGCGCTCAACCCAGC CGC (SEQ ID NO: 1318)	TATGAAGCACTAAGGGCGTcaacttgtgGCGTGGCGCTCAACC	TCACGACGCAGAACTTTACGCAGCGTTAGCTCTTAAACGCGGTG		PARP1_-_226595726 23-P1P2	PARP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAATCTATCAGGGGAAC GG (SEQ ID NO: 1319)	TATGAAGCACTAAGGGCGTcaacttgtgGGCAATCTATCAG	TCACGACGCAGACGTTTACGCAGCTTAGCTCTTAAACCCGTTCCC		PARP1_+_22685771 23-P1P2	PARP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGTGCGGGCGTGTT CGG (SEQ ID NO: 1320)	TATGAACCACTAAGGGCGTcaacttgtgGGCGGGTGCGGGCG	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCCGAACA		PARP1_+_22685747 23-P1P2	PARP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGGCGCGGGCAGCGA GCAC (SEQ ID NO: 1321)	TATGAACCACTAAGGGCGTcaacttgtgGCGCGGCGCGGGC	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGTGCTCC		PARP2_-_20811928 23-P1P2	PARP2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGACGCAGCAGCAGCG CGG (SEQ ID NO: 1322)	TATGAAGCACTAAGGGCGTcaacttgtgCGGCGACGCAGCA	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCCGG		PARP2_-_20811834 23-P1P2	PARP2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGTGCGACGCAGCAC CGA (SEQ ID NO: 1323)	TATGAACCACTAAGGGCGTcaacttgtgCGACGTGCGACGC	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCTTC		PARP2_-_20811833 23-P1P2	PARP2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTCGAGCGCGACGCAGTC CCA (SEQ ID NO: 1324)	TATGAAGCACTAAGGGCGTcaacttgtgAGTCGAGCGCGAC	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTGGGAGT		PARP3_+_51976481 23-P1P2	PARP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTACGCCCTCAAGGACA CAC (SEQ ID NO: 1325)	TATGAAGCACTAAGGGCGTcaacttgtgGCCTACGCCCTCAAG	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGTGTCTC		PARP3_-_51977362 23-P1P2	PARP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCACATGAGCAGCGC CTA (SEQ ID NO: 1326)	TATGAACCACTAAGGGCGTcaacttgtgGGGCCACATGAGC	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTAGCGCT		PARP3_-_51976403 23-P1P2	PARP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na

GAGGTTTGAACTGTATTCAG (SEQ ID NO: 1327)	TATGAACCACTAAGGCGTccaccttgttgGAGGTTTGAACTGTATTCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2900)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGAATACAGTTCAAACCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4473)	1	PARPBP_-_102513970.23-P1P2	PARPBP	P1P2	na	na	na	na	na	na	na	na	na
GCGCGGGAGGGCATCCCGTT (SEQ ID NO: 1328)	TATGAACCACTAAGGCGTccaccttgttgGCGCGGGAGGGCATCCCGTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2901)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAACGGGATGCCCTCCCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4474)	1	PARPBP_-_102514014.23-P1P2	PARPBP	P1P2	na	na	na	na	na	na	na	na	na
GGCGAGACTTACGTGATTAG (SEQ ID NO: 1329)	TATGAACCACTAAGGCGTccaccttgttgGGCGAGACTTACGTGATTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2902)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTAATCACGTAAGTCTCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4475)	1	PARPBP_+_102514053.23-P1P2	PARPBP	P1P2	na	na	na	na	na	na	na	na	na
GCCTGAGGAGATGTTCAGGG (SEQ ID NO: 1330)	TATGAACCACTAAGGCGTccaccttgttgGCCTGAGGAGATGTTCAGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2903)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCTGAACATCTCCTCAGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4476)	1	PAXIP1_+_154794593.23-P1P2	PAXIP1	P1P2	na	na	na	na	na	na	na	na	na
GCGGGCCGCCGCGATCATGT (SEQ ID NO: 1331)	TATGAACCACTAAGGCGTccaccttgttgGCGGGCCGCCGCGATCATGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2904)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACATGATCGCGGCGGCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4477)	1	PAXIP1_+_154794632.23-P1P2	PAXIP1	P1P2	na	na	na	na	na	na	na	na	na
GGCCTGGTCCGACATGATCG (SEQ ID NO: 1332)	TATGAACCACTAAGGCGTccaccttgttgGGCCTGGTCCGACATGATCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2905)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGATCATGTCGGACCAGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4478)	1	PAXIP1_-_154794646.23-P1P2	PAXIP1	P1P2	na	na	na	na	na	na	na	na	na
GGCGGCACCGGGAGGGGTAA (SEQ ID NO: 1333)	TATGAACCACTAAGGCGTccaccttgttgGgcggcaccgggaggggtaagtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2906)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTACCCCTCCCGGTGCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4479)	1	PBRM1_+_52719898.23-P1P2	PBRM1	P1P2	na	na	na	na	na	na	na	na	na
GGCGGCGGGTGCCGAAGCGA (SEQ ID NO: 1334)	TATGAACCACTAAGGCGTccaccttgttgGgcggcgggtgccgaagcgagtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2907)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCTTCGGCACCCGCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4480)	1	PBRM1_+_52719868.23-P1P2	PBRM1	P1P2	na	na	na	na	na	na	na	na	na
GTTTGCGGCGGCACCGGGAG (SEQ ID NO: 1335)	TATGAACCACTAAGGCGTccaccttgttgGTTTgcggcggcaccgggaggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2908)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCCGGTGCCGCCGCAAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4481)	1	PBRM1_+_52719904.23-P1P2	PBRM1	P1P2	na	na	na	na	na	na	na	na	na
GGTCAGGCGAGTCCACGTGA (SEQ ID NO: 1336)	TATGAACCACTAAGGCGTccaccttgttgGGTCAGGCGAGTCCACGTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2909)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCACGTGGACTCGCCTGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4482)	1	PCGF1_+_74734467.23-P1P2	PCGF1	P1P2	na	na	na	na	na	na	na	na	na
GTAAGATGGCGTCTCCTCAG (SEQ ID NO: 1337)	TATGAACCACTAAGGCGTccaccttgttgGTAAGATGGCGTCTCCTCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2910)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGAGGAGACGCCATCTTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4483)	1	PCGF1_+_74734777.23-P1P2	PCGF1	P1P2	na	na	na	na	na	na	na	na	na
GTCGCAATCTGGCCCCCCTG (SEQ ID NO: 1338)	TATGAACCACTAAGGCGTccaccttgttgGTCGCAATCTGGCCCCCCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2911)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGGGGGGCCAGATTGCGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4484)	1	PCGF1_-_74734785.23-P1P2	PCGF1	P1P2	na	na	na	na	na	na	na	na	na
GGAGAAACCTACGGTAAGAA (SEQ ID NO: 1339)	TATGAACCACTAAGGCGTccaccttgttgGGAGAAACCTACGGTAAGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2912)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCTTACCGTAGGTTTCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4485)	1	PCGF2_-_36904575.23-P1	PCGF2	P1	na	na	na	na	na	na	na	na	na
GGCCGCCCGGGCGAGCGACA (SEQ ID NO: 1340)	TATGAACCACTAAGGCGTccaccttgttgGGCCGCCCGGGCGAGCGACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2913)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTCGCTCGCCCGGGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4486)	1	PCGF2_+_36904525.23-P1	PCGF2	P1	na	na	na	na	na	na	na	na	na
GGCGGGAGGGGAGAAACCTA (SEQ ID NO: 1341)	TATGAACCACTAAGGCGTccaccttgttgGGCGGGAGGGGAGAAACCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2914)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGGTTTCTCCCCTCCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4487)	1	PCGF2_-_36904566.23-P1	PCGF2	P1	na	na	na	na	na	na	na	na	na
GCCGCGAAGCTGGGGGCCTC (SEQ ID NO: 1342)	TATGAACCACTAAGGCGTccaccttgttgGCCGCGAAGCTGGGGGCCTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2915)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGGCCCCCAGCTTCGCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4488)	1	PCGF2_-_36905986.23-P2	PCGF2	P2	na	na	na	na	na	na	na	na	na
GCTCCTGGTCGCCGGCCAGA (SEQ ID NO: 1343)	TATGAACCACTAAGGCGTccaccttgttgGCTCCTGGTCGCCGGCCAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2916)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTGGCCGGCGACCAGGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4489)	1	PCGF2_+_36905999.23-P2	PCGF2	P2	na	na	na	na	na	na	na	na	na

GTCGACGCTGGGGTCACT CGG (SEQ ID NO: 1344)	TATGAACCACTAAGGCGTcaacttgtggTCGAGCTGGGGT	TCACCACGCAGACTTTACGCAGCTTACGTCTTAAACCCGAGTG ACCCCAGCTCGAGCCAACAAAGGTCGCAGCGCTTAGTGGTTCATA (SEQ ID NO: 4490)	1	PCGF2_+_3 6806836.23-P2	PCGF2	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCTGGGTCCCCCCC AAG (SEQ ID NO: 1345)	TATGAACCACTAAGGCGTcaacttgtgGCCTCTGGGTCCC	TCACCAGGCAGAGCTTACGCAGCTTAAGCCTTTAAACCTTGGGG GGGACCCCAGCAGGCCAACAAAGGTTGGAGCGCTTAGTCGTTCAT A (SEQ ID NO: 4491)	1	PCGF3_+_6 99865.23-P1P2	PCGF3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCCGGTCCCCCCCCAAG AGG (SEQ ID NO: 1346)	TATGAACCACTAAGGCGTcaacttgtgCTCCGGTCCCCC	TCACCAGCAGAGCTTACGCAGCTTAAGCTTTAAACCCCTCTTGG GGGGGACCCAGCGCAACAAAGGTTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4492)	1	PCGF3_+_6 99862.23-P1P2	PCGF3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGAGAAAAAAAGCCC GTC (SEQ ID NO: 1347)	TATGAACCACTAAGGCGTcaacttgtgCGAGGAGAAAAAA	TCACCAGCAGAGCTTACGCAGCTTAAGCGTTAAACGAGGAGCC TTTTTCTCTCCTCCAACAAAGGTTGGAGCGCTTAGTCGTTCATA (SEQ ID NO: 4493)	1	PCGF3_+_6 99944.23-P1P2	PCGF3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACTCGGCTCAGGGCCCG AGA (SEQ ID NO: 1348)	TATGAACCACTAAGGCGTcaacttgtgGACTCGGCTCAGG	TCACCAGCAGAGCTTACGCAGCTTAAGCTGTTAAACTGTCGGG CCTGAGCCGAGTCCAAGAAAGGTTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4494)	1	PCGF5_+_9 2980415.23-ENST0000036126.5	PCGF5	ENST0000036126.5	PCGF5_+_9 2980415.23-ENST0000043648.1	PCGF5	ENST0000005946	na	na	na	na	na	na	na	na	na	na
GGGTCCGGGTGGCGACTTC CAA (SEQ ID NO: 1349)	TATGAACCACTAAGGCGTcaacttgtgGGGTCCGGGTGG	TCACCAGCAGAGCTTACGCAGCTTAAGCTGTTAAACTTGGAGGG CGGCACCGGAACCCCAACAAAGGTTGGAGCGCTTAGTGGTTCAT A (SEQ ID NO: 4495)	2	PCGF5_+_9 2980372.23-ENST0000036126.5	PCGF5	ENST0000036126.5	PCGF5_+_9 2980372.23-ENST0000043648.1	PCGF5	ENST0000005946	na	na	na	na	na	na	na	na	na	na
GGTGGAGTTTGCCCTTGG AGG (SEQ ID NO: 1350)	TATGAACCACTAAGGCGTcaacttgtgGGTGGAGTTTGCC	TCACCAGCAGAGCTTACGCAGCTTAAGCCTTAAACCCTCCAAGG GGCAACTCCACCCAACAAGGTTGGAGCGCTTAGTGGTTCATA (SEQ ID NO: 4496)	1	PCGF5_-_92980364.23-ENST0000036126.5	PCGF5	ENST0000036126.5	PCGF5_-_92980364.23-ENST0000005946.1	PCGF5	ENST0000005946	na	na	na	na	na	na	na	na	na	na
GAGTGTGGGTCCAGAGA TCA (SEQ ID NO: 1351)	TATGAACCACTAAGGCGTcaacttgtgGAGTGTGGGTCCA	TCACCAGCAGAGCTTACGCAGCTTAAGCTGTTAAACTGATCTCT GGAGCCCAGACTCCAAACAAAGGTCGCGCTTAGTGGTTCATA (SEQ ID NO: 4497)	1	PCGF5_+_9 2979941.23-ENST0000043648.1	PCGF5	ENST0000043648.1	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGATAAGGTGCAAATC AA (SEQ ID NO: 1352)	TATGAACCACTAAGGCGTcaacttgtgGCAGATAAGGTGC	TCACCAGCAGAGCTTACGCAGCTTAAGCGTTAAACTTGATTTG CACGTTATCTGCCAACAAAGGTCGCAGCGCTTAGTGGTTCATA (SEQ ID NO: 4498)	1	PCGF5_-_92980370.2-ENST0000043648.1	PCGF5	ENST0000043648.1	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCATGCGTCGCGCAGACAC ACC (SEQ ID NO: 1353)	TATGAACCACTAAGGCGTcaacttgtgGCATGCGTCGCG	TCACCAGCAGAGCTTACGCAGCTTAAGCGTTAAACGCATGCGT CTCCCGACCATGCCAACAAAGGTCGCAGCGCTTAGTGGTTCATA (SEQ ID NO: 4499)	1	PCGF6_-_105110641.23-P1P2	PCGF6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGGGAGAGACACGAG GCG (SEQ ID NO: 1354)	TATGAACCACTAAGGCGTcaacttgtgGTCGGGAGAGACA	TCACCAGCAGAGCTTACGCAGCTTAAGCGTTAAACGCGCCTCG TGTCTCTCCCGAACCAAGGTCGCGACGCTTAGTGGTTCATA (SEQ ID NO: 4500)	1	PCGF6_-_105110846.23-P1P2	PCGF6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCTCTCCCGACCATGCAT GG (SEQ ID NO: 1355)	TATGAACCACTAAGGCGTcaacttgtgGTCTCTCCCGACCA	TCACCAGCAGAGCTTACGCAGCTTAAGCGTTAAACCCCTCCATG GTCGGGAGAGACACCAACAAAGGTCGCGACGCTTAGTGGTTCATA (SEQ ID NO: 4501)	1	PCGF6_+_1 0511083.23-P1P2	PCGF6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGAGCCTCAGAGCGG AGC (SEQ ID NO: 1356)	TATGAACCACTAAGGCGTcaacttgtgCAGGAGGCCTCAG	TCACCAGCAGAGCTTACGCAGCTTAAGCGTTAAACGCTGCGCT CTGAGGCCTCCTGCCAACAAAGGTCGACGCTTAGTGGTTCATA (SEQ ID NO: 4502)	2	PCNA_-_5100533.23-P1P2	PCNA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGTGGTGACGTCGCA CCG (SEQ ID NO: 1357)	TATGAACCACTAAGGCGTcaacttgtgGCGTGGTGACGT	TCACCAGCAGAGCTTTACGCAGCGTTAAGCGTTAAACGCGTCG CCGCACCACCGCCCAACAAAGGTCGACGCTTAGTGGTTCATA (SEQ ID NO: 4503)	2	PCNA_+_51 005640.23-P1P2	PCNA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGCGAACGTCGCGAC GAC (SEQ ID NO: 1358)	TATGAACCACTAAGGCGTcaacttgtgGGGGCGAACGTC	TCACCAGCAGAGCTTTACGCAGCGTTAAGCGTTCGTCGC CGAGCGTTCCGCCCCAACAAAGGTCGACGCTTAGTGGTTCATA (SEQ ID NO: 4504)	2	PCNA_-_5100552.23-P1P2	PCNA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGCGATTCGATCGCG CCG (SEQ ID NO: 1359)	TATGAACCACTAAGGCGTcaacttgtgGAGGCGATTCGA	TCACCAGCAGAGCTTTACGCAGCGTTAAGCGTTAAACGCCGGCGC ATCCAATCGCGCTCCAACAAAGGTCGACGCGTTAGTGGTTCATA (SEQ ID NO: 4505)	1	PDGFB_-_33160374.23-P1P2	PDGFB	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CGAGTAGCGAGTCGGGCA ACC (SEQ ID NO: 1360)	TATGAACCACTAAGGCGTcaacttgtgCGAGTAGCGAGT	TCACCAGCAGAGCTTTACGCAGCGTTAAGCGTTAAACGGTTGCGC GACTCGCTACTCGCCAACAAGGTCGCGACGCTTAGTGGTTCATA (SEQ ID NO: 4506)	1	PDGFB_-_33160727.23-P1P2	PDGFB	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTAGTCGAGTGGGGAAGC CGG (SEQ ID NO: 1361)	TATGAACCACTAAGGCGTcaacctgtggGTAGCGAGTTGGGC	TCAGCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACCCGGGTT	1	PDS68_-_33160730.2 3-P1P2	PDS68	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGCGGTGAGTGCTCGG GGA (SEQ ID NO: 1362)	TATGAACCACTAAGGCGTcaacctgtgGAGCGGTGAGTGAGTGC	TCAGCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACTCCCCGA	1	PER1_+_8058306.23-P1	PER1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGTCTAGAGCGGAGC GGG (SEQ ID NO: 1363)	TATGAACCACTAAGGCGTcaacctgtgGAGGTCTAGAGTG	TCAACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCCGTC	1	PER1_-_8059436.23-P1	PER1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGCAACCGGGTGAGTG CTC (SEQ ID NO: 1364)	TATGAACCACTAAGGCGTcaacctgtgCGCGCCGAGCGGTG	TCAGCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACGAGCACT	1	PER1_+_8059111.23-P1	PER1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGCAGTCGTGCGACCA GGT (SEQ ID NO: 1365)	TATGAACCACTAAGGCGTcaacctgtgGAGCACTCGTCGCG	TCAGCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACACCTGGT	1	PER1_+_8059649.23-P2	PER1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCCTCGTGAGAGCTTCC ACT (SEQ ID NO: 1366)	TATGAACCACTAAGGCGTcaacctgtgGGCCCGTGAGAGC	TCAACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACAGTGGAA	1	PER1_+_8059707.23-P2	PER1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCGGTGAGAGCTTCC ACC (SEQ ID NO: 1367)	TATGAACCACTAAGGCGTcaacctgtgGGGCCGAGAGAAT	TCAGCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACGGTAGGC	1	PER1_-_8059857.23-P2	PER1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCAACCGGGCGAGCAGA CACT (SEQ ID NO: 1368)	TATGAACCACTAAGGCGTcaacctgtgGCCAACCGGGCGAG	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACAGTTCTGC	1	PHC1_-_9066879.23-P1P2	PHC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGCGGGGAAGGTGCTT CCG (SEQ ID NO: 1369)	TATGAACCACTAAGGCGTcaacctgtgGCTGGGAGGAAGGT	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCGGAAGC	1	PHC1_+_9067071.23-P1P2	PHC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGAACCGCAACCTCTCG GAG (SEQ ID NO: 1370)	TATGAAGCACTAAGGCGTcaacctgtgCGGGCAACCGCAACT	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCTCCGAG	1	PHC1_+_9068302.23-P1P2	PHC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGCGGCGCGCGGAGACA ATG (SEQ ID NO: 1371)	TATGAACCACTAAGGCGTcaacctgtgGAGCGCGCGCGGAG	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCATTGTCT	1	PHC2_-_33815580.23-P1P2	PHC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAATGCGGCGACGTCTG GGA (SEQ ID NO: 1372)	TATGAACCACTAAGGCGTcaacctgtgGCAATGCGGCGCGAG	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTCCCAGAG	1	PHC2_-_33815515.23-P1P2	PHC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGACGCGCTCGAGCGG CCG (SEQ ID NO: 1373)	TATGAAGCACTAAGGCGTcaacctgtgGTCGAGCGCTCGCA	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAAGCCGCGGG	1	PHC2_-_33815498.23-P1P2	PHC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGCGGAGTGAGTGACT AGA (SEQ ID NO: 1374)	TATGAACCACTAAGGCGTcaacctgtgGAGCGGGACTGAGT	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAAGTCTAGTCA	1	PHC3_+_169389473.23-P1P2	PHC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCATGTCGTAACTAGG GG (SEQ ID NO: 1375)	TATGAACCACTAAGGCGTcaacctgtgGCTCATGTCGTAACT	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCCTAGT	1	PHC3_+_169389448.23-P1P2	PHC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCGCGACGCCCATGTTAC TGA (SEQ ID NO: 1376)	TATGAAGCACTAAGGCGTcaacctgtgCGCGCGACGCCCATG	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTCACTAAC	1	PHC3_+_169389513.23-P1P2	PHC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAATGGAGCGAGTGCCCAG CGG (SEQ ID NO: 1377)	TATGAACCACTAAGGCGTcaacctgtgGAATGGAGCGAGTG	GCCTCCTCCACTTCCAACAGAGGTCGAGCAGCTCTTAGTGGTTCATA	1	PHF10_-_17012431523-P1P2	PHF10	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GAGTACAACGGCTCGCGC TTG (SEQ ID NO: 1378)	TATGAACCACTAAGGCGTcaacctgtgGagtacaacgcgggtcgt	TCACCAGCACAGACTTTAGCGCAGCTTAGCTCTTAAACCACGGCC (SEQ ID NO: 4524)	CGCGGTTGTACTCCACAACAAGGTGGACCCGCCTTACTGGTTCATA	1	PHF10_+_17 0124573.23-P1P2	PHF10	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGCGCGCGCTGAGAC CGG (SEQ ID NO: 1379)	TATGAACCACTAAGGCGTcaacctgtgGCGCGCGCGCGCTG AGACCGCGGgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGCAGACTTTAGCGCAGCTTAGCTCTTAAACCCGGTCT (SEQ ID NO: 4525)	CACGCGCGCGCGCCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4525)	1	PHF10_+_17 0123802.23-P1P2	PHF10	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CATAGGAAGTCCGGAAGC TAG (SEQ ID NO: 1380)	TATGAACCACTAAGGCGTcaacctgtgAATACGGCAGTCGCG GAAGTTAGgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACTTTAGCGCAGCTTAGCTCTTAAACCTACCTCC (SEQ ID NO: 4526)	GGACTCCCTATCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIAS1_-_68946973.2-P1P2	PIAS1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGCAACAGTCCGCGAAC TAA (SEQ ID NO: 1381)	TATGAACCACTAAGGCGTcaacctgtgCGCGCAACAAGTCGCG GAACTAAgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACTTACGTCC (SEQ ID NO: 4527)	GCACTGTCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIAS1_-_68946968.2-P1P2	PIAS1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTGACAGAGCGCAAGA TGG (SEQ ID NO: 1382)	TATGAACCACTAAGGCGTcaacctgtgGGGTGACAGAGCGC AAGATGGgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACGTTTAGCGCAGCTTAGCTCTTAAACCCATCTTG (SEQ ID NO: 4528)	CGTCTGTCACGCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIAS1_-_68946970.2-P1P2	PIAS1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCCACCGCTCACTCA CTT (SEQ ID NO: 1383)	TATGAACCACTAAGGCGTcaacctgtgGCGCCACCGGTCA CTCACTTgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACAGACTTTAGCTCTTAAACAAGTGAGTT (SEQ ID NO: 4529)	GACGCGTTGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIAS4_+_40 07760.23-P1P2	PIAS4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGACGCTGGTGACCA AGA (SEQ ID NO: 1384)	TATGAACCACTAAGGCGTcaacctgtgCGGCGACGCTGGTG ACCAACGgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACGTTTAGCGCAGCTTAGCTCTTAAACTCTTGGTC (SEQ ID NO: 4530)	ACCACGCGTCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIAS4_-_40077761.23-P1P2	PIAS4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCACCAGCGTCCCCCGG CGC (SEQ ID NO: 1385)	TATGAACCACTAAGGCGTcaacctgtgGTCACCAGCGCTCC CCCCGCCGgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACGTTTAGCGCAGCTTAGCTCTTAAAGCGCGGGG (SEQ ID NO: 4531)	CGGACGCGTCGTGACCACAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIAS4_+_40 07731.23-P1P2	PIAS4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTGTCTGCCCGAAGC CAG (SEQ ID NO: 1386)	TATGAACCACTAAGGCGTcaacctgtgGAGTGTCTGCCCG AACCCAAgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACTTTAGCGCAGCTTAGCTCTTAAACCCTGGGTT (SEQ ID NO: 4532)	CGGGCAGACACTCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIF1_+_35111 7706.23-P1	PIF1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCTTTGCTCCCCAGCT CT (SEQ ID NO: 1387)	TATGAAGCACTAAGGCGTcaacctgtgGCTCTTTGCTCCCC ACCTCTgttaagactaagctgGGTAAAGTTCTGCTGGTGA	TCAGCAGCAGACGTTTAGCGCAGCTTAGCTCTTAAACAGAGTGTC (SEQ ID NO: 4533)	GAGAGCGAAAGAGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4533)	1	PIF1_-_65117773.2-P1	PIF1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCTCGCCCGAACCCAG AGG (SEQ ID NO: 1388)	TATGAACCACTAAGGCGTcaacctgtgGGTCTCGCCCGAAC CCAGAGGgttaagactaagctgGGTAAAGTTCTGCTGGTGA	TCACCAGCAGACGTTTAGCGCAGCTTAGCTCTTAAACCCTCTGG (SEQ ID NO: 4534)	GTTCGGGCGACGAACCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIF1_+_86111 7763.23-P1	PIF1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGTGAGGACGGCGAAA GAA (SEQ ID NO: 1389)	TATGAACCACTAAGGCGTcaacctgtgGCAGTGAGGACGG GAAAAGAAgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACAGACTTTAGCTCTTAAACTTCTTTCC (SEQ ID NO: 4535)	CGTCCTCACTCGCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIF1_+_35111 5366.23-P2	PIF1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCGAAGTGTGCGCAGG AAG (SEQ ID NO: 1390)	TTATGAAGCACTAAGGCGTcaacctgtgCGCGCGAATGTGCG TCAGGAAGgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCAGCAGCAGACGTTTAGCGCAGCTTAGCTCTTAAACCTTCCTCC (SEQ ID NO: 4536)	GCACAGATTGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIF1_-_55118192.2-P2	PIF1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCCGAGGCTAGCACA GTG (SEQ ID NO: 1391)	TATGAACCACTAAGGCGTcaacctgtgGGGGCCGAGGCTAG CACAGTGgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACAGACTTTAGCGCAGCTTAGCTCTTAAACGCACTGTG (SEQ ID NO: 4537)	CTAGGCTCGGCCGCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PIF1_+_35111 5970.23-P2	PIF1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CGAAGGTCTGCAGCGCGCAG TCT (SEQ ID NO: 1392)	TTATGAACCACTAAGGCGTcaacctgtgCGAAGGTCTGCAG CGCAGCTTgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGCCTCGCGAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4538)	GCTGCAGACCTCGCGAACAAGGTGGACGCCTTAGTGGTTCATA	1	PLK1_-_23692248.2-P1P2	PLK1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGTGCTCTGCAGTGACT GCA (SEQ ID NO: 1393)	TATGAACCACTAAGGCGTcaacctgtgGTGAGTGCTCGCAGT GACTGCAgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACGTTTAGCGCAGCTTAGCTCTTAAACTGCACGTCA (SEQ ID NO: 4539)	CTCGCAGCACTCCCAACCAAGGTGGACGCCTTAGTGGTTCATA	1	PLK1_-_23690276.2-P1P2	PLK1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGGCTCTGCTGGGA TCG (SEQ ID NO: 1394)	TATGAACCACTAAGGCGTcaacctgtgGGAAGGCTCTGCT GGGATCGgttaagactaagctgGGTAAAGTCTGCTGGTGA	TCACCAGCAGACAGACTTTAGCGCAGCTTAGCTCTTAAACCGATCCCE (SEQ ID NO: 4540)	AGCCAGAGACGTCTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA	1	PLK1_-_23690230.2-P1P2	PLK1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCGGACCAGGAACGAGTA GAG (SEQ ID NO: 1395)	TATGAACCACTAAGGCTcaacctgtgGCGGAACCAGCAG GAGTGTCGTCCGCCAACAAGGTGTGGACGCCTTAGTCGTTCATA (SEQ ID NO: 2958)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACCTCTACTC 84937823-P1P2	PIP2	na	na	na	na	PMS1_+_190849378 23-P1P2	PIP2	na	na	na	na	na	na	na	na	na	na
GTCCGGGCGCGCAAACGCT ACTA (SEQ ID NO: 1396)	TATGAACCACTAAGGCTcaacctgtgGTCCGGGCGCGCGAA TCGGCTATAgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2959)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACTATAGCGT TCGGGCCCGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4542)	PIP2	na	na	na	na	PMS1_-_190649415 23-P1P2	PIP2	na	na	na	na	na	na	na	na	na	na
GTGCGCATCGGGTGAGA CGC (SEQ ID NO: 1397)	TATGAACCACTAAGGCTcaacctgtgGTCGCGATCCGCGT GAGAGCGCgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2970)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACGCGGTCGTC ACCCGATGCGACGAAGAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4543)	PIP2	na	na	na	na	PMS1_-_196649259 23-P1P2	PIP2	na	na	na	na	na	na	na	na	na	na
GCCAATGGGAGTTCACGC AGG (SEQ ID NO: 1398)	TATGAACCACTAAGGCTcaacctgtgGCCAATTCGCGAGTT ACTCCCATTgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2971)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACCCTCGTGA 8713 23-P1P2	PIP2	na	na	na	na	PMS2_+_80487133-P1P2	PMS2	na	na	na	na	na	na	na	na	na	na
GGCCTCAAGGCGCACCC AAG (SEQ ID NO: 1399)	TATGAACCACTAAGGCTcaacctgtgGGCCTCAAGGCGC ACCCAAGgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2972)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACTGGATGC GCGCCTTGAGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4545)	PIP2	na	na	na	na	PMS2_-_6049462 23-P1P2	PMS2	na	na	na	na	na	na	na	na	na	na
GGGATCGGGTGTCGTGCAT GCA (SEQ ID NO: 1400)	TATGAACCACTAAGGCTcaacctgtgGGGATCGGGTCGT GCATCCAgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2973)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACTGGATGC AACACCCGATCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4546)	PIP2	na	na	na	na	PMS2_+_80486463 23-P1P2	PMS2	na	na	na	na	na	na	na	na	na	na
GCCACTGCGGCACCCAG GAT (SEQ ID NO: 1401)	TATGAACCACTAAGGCTcaacctgtgGCCACTGCGGCAC CCAGGATgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2974)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACTATCTCTGG GTCCCGCCTGGCCTGGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4547)	PIP2	na	na	na	na	PMKP_+_53377345.23-P1P2	PMKP	na	na	na	na	na	na	na	na	na	na
GGTTGCAGGCGGGGCCAC CTC (SEQ ID NO: 1402)	TATGAACCACTAAGGCTcaacctgtgGGTTGCAGGCGGG GCCACCTCgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2975)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACGAAGGTGC CCCGCCTGCGAACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4548)	PIP2	na	na	na	na	PMKP_+_53370774.23-P1P2	PMKP	na	na	na	na	na	na	na	na	na	na
GTCACGGCCACTTCCGGA CCA (SEQ ID NO: 1403)	TATGAACCACTAAGGCTcaacctgtgGTCACGGCCACTTC AAAGTCGGCGGTCGAGCGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 2976)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACTGGTCGTG 503707852-P1P2	PIP2	na	na	na	na	PMKP_-_503707852 3-P1P2	PMKP	na	na	na	na	na	na	na	na	na	na
GCCTTGGTCGTCCCCGG CGG (SEQ ID NO: 1404)	TATGAAGCACTAAGGCTcaacctgtgGCCTTGGTCGTCC CGGGAGCGgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2977)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACCCGGCCGG GGACCACCAAGGCCACAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4550)	PIP2	na	na	na	na	POLA1_-_247122472-P1P2	POLA1	na	na	na	na	na	na	na	na	na	na
GCGACTGTGTGGAGACTCT CAG (SEQ ID NO: 1405)	TATGAACCACTAAGGCTcaacctgtgGCGACTGTGTGGAGA CTCTCAGgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2978)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACCTGAGAG TCTCCACAGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4551)	PIP2	na	na	na	na	POLA1_+_247123452-P1P2	POLA1	na	na	na	na	na	na	na	na	na	na
GCGCAGACCTTGGTCGT CCC (SEQ ID NO: 1406)	TATGAACCACTAAGGCTcaacctgtgGCGCAGACCTTGG TCGTCCCgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2979)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACGCGGGACGA CCAAGGTCTGGTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4552)	PIP2	na	na	na	na	POLA1_-_247122412-P1P2	POLA1	na	na	na	na	na	na	na	na	na	na
GAGCGGAGCAGAAATCTC GAG (SEQ ID NO: 1407)	TATGAAGCACTAAGGCTcaacctgtgGAGCGGAGCAGAAA TCTCTCGAGgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2980)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACTCCGAGAT TTCTTGTCCCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4553)	PIP2	na	na	na	na	POLA2_+_85023603 23-P1P2	POLA2	na	na	na	na	na	na	na	na	na	na
GGAAGGCAGTGACGTTCTC ACC (SEQ ID NO: 1408)	TATGAACCACTAAGGCTcaacctgtgGGAAGGCAGTGAC GTTCTCACCgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2981)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACGGGTCGAGA AGTTCTCCTGCTTCCCAAGAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4554)	PIP2	na	na	na	na	POLA2_-_85029485 23-P1P2	POLA2	na	na	na	na	na	na	na	na	na	na
CGTTGGCGTTTCCTGGGAATG AGC (SEQ ID NO: 1409)	TATGAACCACTAAGGCTcaacctgtgCGTTGGCGTTTCC TGGGAATGgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2982)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACGGTCGATTC CGCCAACCGCCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4555)	PIP2	na	na	na	na	POLA2_-_85028943 23-P1P2	POLA2	na	na	na	na	na	na	na	na	na	na
GCCGCTCTCCAGCGTTGAA GG (SEQ ID NO: 1410)	TATGAACCACTAAGGCTcaacctgtgGCGCTCTCCAGCGT TGAAGGgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2983)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACCCTTCAAG CTTGGGAGAGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4556)	PIP2	na	na	na	na	POLB_+_42196071 23-P1P2	POLB	na	na	na	na	na	na	na	na	na	na
GGCGCCGCGCCCCCCAT CGG (SEQ ID NO: 1411)	TATGAACCACTAAGGCTcaacctgtgGGCGCCGCGCCCC CCATCGGgttaagagctaagctgtGGTAAAGTCTGGTGGTGA (SEQ ID NO: 2984)	TCACCACGCAGACTTTAGCGCAGCTTAGCTCTTAAACCCGATGGG GGGGCGCGGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4557)	PIP2	na	na	na	na	POLB_-_42196005 23-P1P2	POLB	na	na	na	na	na	na	na	na	na	na

GGGTACCAGGACTTGGAGAC (SEQ ID NO: 1412)	TATGAACCACTAAGGCGTccaccttgttgGGGTACCAGGACTTGGAGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2985)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCTCCAAGTCCTGGTACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4558)	1		POLB_+_42196051.23-P1P2	POLB	P1P2	na	na	na	na	na	na	na	na	na
GCGGGAAACGCTGTTTGAAG (SEQ ID NO: 1413)	TATGAACCACTAAGGCGTccaccttgttgGCGGGAAACGCTGTTTGAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2986)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTCAAACAGCGTTTCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4559)	1		POLD1_-_50887646.23-P1P2	POLD1	P1P2	na	na	na	na	na	na	na	na	na
GTTGAAGCGGGTGAGTAGAG (SEQ ID NO: 1414)	TATGAACCACTAAGGCGTccaccttgttgGTTGAAGCGGGTGAGTAGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2987)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCTACTCACCCGCTTCAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4560)	1		POLD1_-_50887659.23-P1P2	POLD1	P1P2	na	na	na	na	na	na	na	na	na
GTTGCGCGCGCGGGAGTCAG (SEQ ID NO: 1415)	TATGAACCACTAAGGCGTccaccttgttgGTTGCGCGCGCGGGAGTCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2988)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGACTCCCGCGCGCGCAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4561)	1		POLD1_-_50887603.23-P1P2	POLD1	P1P2	na	na	na	na	na	na	na	na	na
GCAGCCAGGCCGCGCTACTC (SEQ ID NO: 1416)	TATGAACCACTAAGGCGTccaccttgttgGCAGCCAGGCCGCGCTACTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2989)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGTAGCGCGGCCTGGCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4562)	1		POLD2_-_44162935.23-P1P2	POLD2	P1P2	na	na	na	na	na	na	na	na	na
GGCGCGGTGAGTGGGGTCCA (SEQ ID NO: 1417)	TATGAACCACTAAGGCGTccaccttgttgGGCGCGGTGAGTGGGGTCCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2990)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGACCCCACTCACCGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4563)	1		POLD2_+_44163092.23-P1P2	POLD2	P1P2	na	na	na	na	na	na	na	na	na
GGCTCGAACGTGCTTGATGG (SEQ ID NO: 1418)	TATGAACCACTAAGGCGTccaccttgttgGGCTCGAACGTGCTTGATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2991)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATCAAGCACGTTCGAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4564)	1		POLD2_-_44162910.23-P1P2	POLD2	P1P2	na	na	na	na	na	na	na	na	na
GAAGCTGGTCCGCCATGGTG (SEQ ID NO: 1419)	TATGAACCACTAAGGCGTccaccttgttgGAAGCTGGTCCGCCATGGTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2992)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCACCATGGCGGACCAGCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4565)	2		POLD3_+_74303696.23-P1P2	POLD3	P1P2	na	na	na	na	na	na	na	na	na
GCAACGGAGGCGACGCACTA (SEQ ID NO: 1420)	TATGAACCACTAAGGCGTccaccttgttgGCAACGGAGGCGACGCACTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2993)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGTGCGTCGCCTCCGTTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4566)	2		POLD3_+_74303572.23-P1P2	POLD3	P1P2	na	na	na	na	na	na	na	na	na
GGAGCAAAGACGTTTCCCGC (SEQ ID NO: 1421)	TATGAACCACTAAGGCGTccaccttgttgGGAGCAAAGACGTTTCCCGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2994)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGGGAAACGTCTTTGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4567)	2		POLD3_-_74303643.23-P1P2	POLD3	P1P2	na	na	na	na	na	na	na	na	na
GAGAGGAGAGACCGGCCAGT (SEQ ID NO: 1422)	TATGAACCACTAAGGCGTccaccttgttgGAGAGGAGAGACCGGCCAGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2995)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACTGGCCGGTCTCTCCTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4568)	1		POLD4_-_67120933.23-P1P2	POLD4	P1P2	na	na	na	na	na	na	na	na	na
GCACCCGAGCTAGGTCTCTG (SEQ ID NO: 1423)	TATGAACCACTAAGGCGTccaccttgttgGCACCCGAGCTAGGTCTCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2996)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGAGACCTAGCTCGGGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4569)	1		POLD4_+_67120763.23-P1P2	POLD4	P1P2	na	na	na	na	na	na	na	na	na
GCTAGAGGAGGCCCCTAGAG (SEQ ID NO: 1424)	TATGAACCACTAAGGCGTccaccttgttgGCTAGAGGAGGCCCCTAGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2997)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCTAGGGGCCTCCTCTAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4570)	1		POLD4_+_67120706.23-P1P2	POLD4	P1P2	na	na	na	na	na	na	na	na	na
GAAGATGGCGGACATCTCCC (SEQ ID NO: 1425)	TATGAACCACTAAGGCGTccaccttgttgGAAGATGGCGGACATCTCCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2998)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGAGATGTCCGCCATCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4571)	1		POLDIP3_+_43010844.23-P1P2	POLDIP3	P1P2	na	na	na	na	na	na	na	na	na
GCTCATCAGGAAGCGCGGGG (SEQ ID NO: 1426)	TATGAACCACTAAGGCGTccaccttgttgGCTCATCAGGAAGCGCGGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 2999)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCCGCGCTTCCTGATGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4572)	1		POLDIP3_+_43010817.23-P1P2	POLDIP3	P1P2	na	na	na	na	na	na	na	na	na
GCTCCCTGGACGAACTCATC (SEQ ID NO: 1427)	TATGAACCACTAAGGCGTccaccttgttgGCTCCCTGGACGAACTCATCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3000)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGATGAGTTCGTCCAGGGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4573)	1		POLDIP3_+_43010830.23-P1P2	POLDIP3	P1P2	na	na	na	na	na	na	na	na	na
GCCCCGAGCCCCTCCGTCGG (SEQ ID NO: 1428)	TATGAACCACTAAGGCGTccaccttgttgGCCCCGAGCCCCTCCGTCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3001)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGACGGAGGGGCTCGGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4574)	1		POLE_-_133263686.23-P1P2	POLE	P1P2	na	na	na	na	na	na	na	na	na

GCTCACGAGACATGGAGC CGT (SEQ ID NO: 1429)	TATGAACCACTAAGGCGTcaacctgtgGCTCAGGAGACATG (SEQ ID NO: 3022)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACACGGCTC CATGTCTCTGAGCCAACAAGGTCGGACCGCCTTAGTGCGTTCATA (SEQ ID NO: 4575)	POLE_-_133283911.23-P1P2	POLE	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCTCTCGCGAGTCTG GGT (SEQ ID NO: 1430)	TATGAACCACTAAGGCGTcaacctgtgGGGCTCTCGCGAG TCTCGGGTgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3023)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACACCCAGA CTCGCGAGAGAGCCCCAAGCAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4576)	POLE_+_133259345.23-P1P2	POLE	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCAAATATGGCGCGC GAG (SEQ ID NO: 1431)	TATGAACCACTAAGGCGTcaacctgtgGCGCAAATATGGC GCGCGAGAgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3024)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACCTCCGCGC GCGCATATTTGCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4577)	POLE2_+_65415466.23-P1P2	POLE2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCAGCTTCAAGGCGC AGA (SEQ ID NO: 1432)	TATGAACCACTAAGGCGTcaacctgtgCGCCAGCTTCAAGCT TTCAAGCTGCCCAAGAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4578)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACTTCCAGCT TCCAAGCTGCCCAAGAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4578)	POLE2_+_65415818.23-P1P2	POLE2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGGCTCGGGACAGCTT GAA (SEQ ID NO: 1433)	TATGAACCACTAAGGCGTcaacctgtgGTGGGCTCGGGCA CCTTGAAgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3026)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACTTCAAGCT GCCCGAGTCCACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4579)	POLE2_+_65415426.23-P1P2	POLE2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCTGAAGCGAGAGCGT AAG (SEQ ID NO: 1434)	TATGAACCACTAAGGCGTcaacctgtgGCCCTGAAGCGAG GGCGAAGgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3027)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACCTTCGCCT CCGCTTCAGGGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4580)	POLE3_+_1176895.23-P1P2	POLE3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTAGCGACCGCGGGAA GTC (SEQ ID NO: 1435)	TATGAACCACTAAGGCGTcaacctgtgCGTAGCGACCGCGG GCGTCGCTACGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3028)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACGACGCTCCC GCGTCGCTACGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4581)	POLE3_+_1176928.23-P1P2	POLE3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGAGCGAACGGGGCA GGA (SEQ ID NO: 1436)	TATGAACCACTAAGGCGTcaacctgtgGGGCGGAGCGAAC GGCGAGGAgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3029)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACTCCTCGCCC GTTCGTCGCTACGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4582)	POLE3_+_1176946.23-P1P2	POLE3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGCGCTACACAGAGG GTG (SEQ ID NO: 1437)	TATGAACCACTAAGGCGTcaacctgtgGCGCGCGCTACAC GTAGGCGCGCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3030)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACCACCCTGTT CATCCGCGCGCGGTTGTTCTGCGGCTTAGTGGTTCATA (SEQ ID NO: 4583)	POLE4_-_75185725.23-P1P2	POLE4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAGCGACGTGCCGCCG TAG (SEQ ID NO: 1438)	TATGAAGCACTAAGGCGTcaacctgtgCGCGACGCGCACTTGC CGACGTAGgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3031)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACCTACGCGC CGACGTAGCGCTGCCCAACAAGGTCGACGCCTTAGTGCGTTCATA (SEQ ID NO: 4584)	POLE4_-_75195697.23-P1P2	POLE4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGCGCGGAACGCCCCC GAG (SEQ ID NO: 1439)	TATGAACCACTAAGGCGTcaacctgtgGGGAAGCGCGGAA CCCCCGAGgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3032)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACCTCGCGCGG CGTCCCCGCTTCCCCCAACAAGGTCGACGCCTTAGTGCGTTCATA (SEQ ID NO: 4585)	POLE4_-_75186948.23-P1P2	POLE4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCGCTCCCTTGACTG GAG (SEQ ID NO: 1440)	TATGAACCACTAAGGCGTcaacctgtgGCCTCGCTCCCTG ACTTGAGgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3033)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACCTCCAGTG AGGGGAGCGAGGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4586)	POLG_-_89877944.23-P1P2	POLG	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCTGCAGCACGGGAAG TTG (SEQ ID NO: 1441)	TATGAAGCACTAAGGCGTcaacctgtgCGGGGTCGCAGCAGG GAAGTTTGgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3034)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACCAACTTCC CGTCTGGCACCCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4587)	POLG_+_898780037.23-P1P2	POLG	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTAGCGTGTGTGGCCTCC ACC (SEQ ID NO: 1442)	TATGAACCACTAAGGCGTcaacctgtgGTAGCGTGTGTGG CCTCCACCgttaagagtaagtgtgCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3035)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACGGGTGCGAG GGCACCACCCTCCACCAACAAGGTCGACGCCTTAGTGCGTTCATA (SEQ ID NO: 4588)	POLG_-_89878816.23-P1P2	POLG	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATTCTGTTAACGGTATGT GG (SEQ ID NO: 1443)	TATGAACCACTAAGGCGTcaacctgtgGATTCTGTTAACGG TAGTCGgttaagagtaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3036)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACCGCTACGG GTTAACATCCAACAAGGTCGACGCCTTAGTGCGTTCATA (SEQ ID NO: 4589)	POLG2_+_8249314023.23-P1P2	POLG2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCTCCGCATTCTGTT AA (SEQ ID NO: 1444)	TATGAACCACTAAGGCGTcaacctgtgGCCTCTCCGCATT TGTTAAgttaagagtaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3037)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACTTAACGAGA ATCCGGAGAGAGCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4590)	POLG2_+_8249314023.23-P1P2	POLG2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGAATTCGTGTTAACGGT AG (SEQ ID NO: 1445)	TATGAACCACTAAGGCGTcaacctgtgGCGGAATTCGTGTTAA CGGTAGgttaagagtaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3038)	TCACCAGCACAGACTTTAGCGAGGTTAGCTCTTAAACGTTACGGTT AACATGAATCCGGCCAACAAGGTCGACGCCTTAGTGCGTTCATA (SEQ ID NO: 4591)	POLG2_+_8249314323.23-P1P2	POLG2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GAGCAGGGATCAGACGAGAG (SEQ ID NO: 1446)	TATGAACCACTAAGGCGTccaccttgttgGAGCAGGGATCAGACGAGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3019)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCTCGTCTGATCCCTGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4592)	1		POLH_+_43543948.23-P1P2	POLH	P1P2	na	na	na	na	na	na	na	na	na
GCAATGCAGCCGCCAGACTG (SEQ ID NO: 1447)	TATGAACCACTAAGGCGTccaccttgttgGCAATGCAGCCGCCAGACTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3020)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGTCTGGCGGCTGCATTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4593)	1		POLH_+_43543917.23-P1P2	POLH	P1P2	na	na	na	na	na	na	na	na	na
GGTCTGGCGGCTGCATTGCT (SEQ ID NO: 1448)	TATGAACCACTAAGGCGTccaccttgttgGGTCTGGCGGCTGCATTGCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3021)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGCAATGCAGCCGCCAGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4594)	1		POLH_-_43543943.23-P1P2	POLH	P1P2	na	na	na	na	na	na	na	na	na
GAAGCTGGGGGTGGAGCCGG (SEQ ID NO: 1449)	TATGAACCACTAAGGCGTccaccttgttgGAAGCTGGGGGTGGAGCCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3022)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGCTCCACCCCCAGCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4595)	1		POLI_-_51795943.23-P1P2	POLI	P1P2	na	na	na	na	na	na	na	na	na
GACGCGGCGGGTACCTCTAG (SEQ ID NO: 1450)	TATGAACCACTAAGGCGTccaccttgttgGACGCGGCGGGTACCTCTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3023)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTAGAGGTACCCGCCGCGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4596)	1		POLI_-_51796203.23-P1P2	POLI	P1P2	na	na	na	na	na	na	na	na	na
GATGGAACTGGCGGACGTGG (SEQ ID NO: 1451)	TATGAACCACTAAGGCGTccaccttgttgGATGGAACTGGCGGACGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3024)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACGTCCGCCAGTTCCATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4597)	1		POLI_-_51796012.23-P1P2	POLI	P1P2	na	na	na	na	na	na	na	na	na
GAGTCGCGATCCTGAGGTAA (SEQ ID NO: 1452)	TATGAACCACTAAGGCGTccaccttgttgGAGTCGCGATCCTGAGGTAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3025)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTACCTCAGGATCGCGACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4598)	1		POLK_-_74807738.23-P1P2	POLK	P1P2	na	na	na	na	na	na	na	na	na
GCGACTGACAAGGCCCCAAG (SEQ ID NO: 1453)	TATGAACCACTAAGGCGTccaccttgttgGCGACTGACAAGGCCCCAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3026)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTGGGGCCTTGTCAGTCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4599)	1		POLK_+_74807767.23-P1P2	POLK	P1P2	na	na	na	na	na	na	na	na	na
GGGACAAGCGATCCCCGCGC (SEQ ID NO: 1454)	TATGAACCACTAAGGCGTccaccttgttgGGGACAAGCGATCCCCGCGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3027)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGCGGGGATCGCTTGTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4600)	1		POLK_+_74807747.23-P1P2	POLK	P1P2	na	na	na	na	na	na	na	na	na
GAAGCTAGTCACCCGGGGGT (SEQ ID NO: 1455)	TATGAACCACTAAGGCGTccaccttgttgGAAGCTAGTCACCCGGGGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3028)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCCCCGGGTGACTAGCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4601)	1		POLL_-_10334794123-P1P2	POLL	P1P2	na	na	na	na	na	na	na	na	na
GAGACTCTTGGGGCCGGGGT (SEQ ID NO: 1456)	TATGAACCACTAAGGCGTccaccttgttgGAGACTCTTGGGGCCGGGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3029)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCCCGGCCCCAAGAGTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4602)	1		POLL_+_103347853.23-P1P2	POLL	P1P2	na	na	na	na	na	na	na	na	na
GGGCCAAGCTAGTCACCCGG (SEQ ID NO: 1457)	TATGAACCACTAAGGCGTccaccttgttgGGGCCAAGCTAGTCACCCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3030)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGGTGACTAGCTTGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4603)	1		POLL_-_103347937.23-P1P2	POLL	P1P2	na	na	na	na	na	na	na	na	na
GAGGGAAACTCCGAGCGAGA (SEQ ID NO: 1458)	TATGAACCACTAAGGCGTccaccttgttgGAGGGAAACTCCGAGCGAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3031)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTCGCTCGGAGTTTCCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4604)	1		POLM_-_44122095.23-P1	POLM	P1	na	na	na	na	na	na	na	na	na
GCAGCGCGGAGCGAACGCAG (SEQ ID NO: 1459)	TATGAACCACTAAGGCGTccaccttgttgGCAGCGCGGAGCGAACGCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3032)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGCGTTCGCTCCGCGCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4605)	1		POLM_-_44122076.23-P1	POLM	P1	na	na	na	na	na	na	na	na	na
GCCCAATGCTCCCCAAACGG (SEQ ID NO: 1460)	TATGAACCACTAAGGCGTccaccttgttgGCCCAATGCTCCCCAAACGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3033)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGTTTGGGGAGCATTGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4606)	1		POLM_+_44122019.23-P1	POLM	P1	na	na	na	na	na	na	na	na	na
GCAAGGAGGGTCAGGAGGCG (SEQ ID NO: 1461)	TATGAACCACTAAGGCGTccaccttgttgGCAAGGAGGGTCAGGAGGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3034)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCTCCTGACCCTCCTTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4607)	1		POLM_+_44113765.23-P2	POLM	P2	na	na	na	na	na	na	na	na	na
GGCAGGCGGCACATCACTCT (SEQ ID NO: 1462)	TATGAACCACTAAGGCGTccaccttgttgGGCAGGCGGCACATCACTCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3035)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGAGTGATGTGCCGCCTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4608)	1		POLM_-_44113758.23-P2	POLM	P2	na	na	na	na	na	na	na	na	na

GTGGCCAGCGGTGGCTCTTG (SEQ ID NO: 1463)	TATGAACCACTAAGGCGTccaccttgttgGTGGCCAGCGGTGGCTCTTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3036)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAAGAGCCACCGCTGGCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4609)	1		POLM_-_44113711.23-P2	POLM	P2	na	na	na	na	na	na	na	na	na
GCGCGCCTCGCACTGCCTCA (SEQ ID NO: 1464)	TATGAACCACTAAGGCGTccaccttgttgGCGCGCCTCGCACTGCCTCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3037)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGAGGCAGTGCGAGGCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4610)	1		POLN_-_2243856.23-P1P2	POLN	P1P2	HAUS3_-_2243856.23-P1P2	HAUS3	P1P2	na	na	na	na	na	na
GGGCCCGCGTCAGTCACCTC (SEQ ID NO: 1465)	TATGAACCACTAAGGCGTccaccttgttgGGGCCCGCGTCAGTCACCTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3038)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGGTGACTGACGCGGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4611)	1		POLN_-_2243782.23-P1P2	POLN	P1P2	HAUS3_-_2243782.23-P1P2	HAUS3	P1P2	na	na	na	na	na	na
GGGCTCCTGCGAGAAGCAAG (SEQ ID NO: 1466)	TATGAACCACTAAGGCGTccaccttgttgGGGCTCCTGCGAGAAGCAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3039)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTGCTTCTCGCAGGAGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4612)	1		POLN_+_2243788.23-P1P2	POLN	P1P2	HAUS3_+_2243788.23-P1P2	HAUS3	P1P2	na	na	na	na	na	na
GCAGGGTTCTCCCGAGAGGG (SEQ ID NO: 1467)	TATGAACCACTAAGGCGTccaccttgttgGCAGGGTTCTCCCGAGAGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3040)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTCTCGGGAGAACCCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4613)	1		POLQ_+_121264769.23-P1P2	POLQ	P1P2	na	na	na	na	na	na	na	na	na
GGGCCAGGGTTCTCCCGAGA (SEQ ID NO: 1468)	TATGAACCACTAAGGCGTccaccttgttgGGGCCAGGGTTCTCCCGAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3041)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTCGGGAGAACCCTGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4614)	1		POLQ_+_121264772.23-P1P2	POLQ	P1P2	na	na	na	na	na	na	na	na	na
GGTCCCAGCGTCCTCCCTCT (SEQ ID NO: 1469)	TATGAACCACTAAGGCGTccaccttgttgGGTCCCAGCGTCCTCCCTCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3042)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGAGGGAGGACGCTGGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4615)	1		POLQ_-_121264780.23-P1P2	POLQ	P1P2	na	na	na	na	na	na	na	na	na
GGCAAGCGCGGAAGGCTGGG (SEQ ID NO: 1470)	TATGAACCACTAAGGCGTccaccttgttgGGCAAGCGCGGAAGGCTGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3043)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCAGCCTTCCGCGCTTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4616)	2		POLR2I_+_36605780.23-P1P2	POLR2I	P1P2	na	na	na	na	na	na	na	na	na
GGTTCTTCCAATAGCGGCCC (SEQ ID NO: 1471)	TATGAACCACTAAGGCGTccaccttgttgGGTTCTTCCAATAGCGGCCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3044)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGCCGCTATTGGAAGAACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4617)	2		POLR2I_+_36606198.23-P1P2	POLR2I	P1P2	na	na	na	na	na	na	na	na	na
GTTCACGGTTCTTCCAATAG (SEQ ID NO: 1472)	TATGAACCACTAAGGCGTccaccttgttgGTTCACGGTTCTTCCAATAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3045)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTATTGGAAGAACCGTGAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4618)	2		POLR2I_+_36606204.23-P1P2	POLR2I	P1P2	na	na	na	na	na	na	na	na	na
GAACATGGCCGCCGCCTGAG (SEQ ID NO: 1473)	TATGAACCACTAAGGCGTccaccttgttgGAACATGGCCGCCGCCTGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3046)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCAGGCGGCGGCCATGTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4619)	1		POLR3E_-_22308778.23-P1P2	POLR3E	P1P2	na	na	na	na	na	na	na	na	na
GCGCCGCCTGAGAGGAGAGC (SEQ ID NO: 1474)	TATGAACCACTAAGGCGTccaccttgttgGCGCCGCCTGAGAGGAGAGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3047)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCTCTCCTCTCAGGCGGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4620)	1		POLR3E_-_22308786.23-P1P2	POLR3E	P1P2	na	na	na	na	na	na	na	na	na
GTACCCGCGGGCTGCAGAGA (SEQ ID NO: 1475)	TATGAACCACTAAGGCGTccaccttgttgGTACCCGCGGGCTGCAGAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3048)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTCTGCAGCCCGCGGGTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4621)	1		POLR3E_+_22308794.23-P1P2	POLR3E	P1P2	na	na	na	na	na	na	na	na	na
GCGCGCGTCCACCGGCTCCC (SEQ ID NO: 1476)	TATGAACCACTAAGGCGTccaccttgttgGcgcgcgtccaccggctcccgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3049)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGAGCCGGTGGACGCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4622)	1		PPP4R1_+_9614544.23-P1P2	PPP4R1	P1P2	na	na	na	na	na	na	na	na	na
GGCCGCCGCCCGGAGAACAG (SEQ ID NO: 1477)	TATGAACCACTAAGGCGTccaccttgttgGGCCGCCGCCCGGAGAACAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3050)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGTTCTCCGGGCGGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4623)	1		PPP4R1_-_9614465.23-P1P2	PPP4R1	P1P2	na	na	na	na	na	na	na	na	na
GGCGCCGCTCCATGTAGCCC (SEQ ID NO: 1478)	TATGAACCACTAAGGCGTccaccttgttgGgcgccgctccatgtagcccgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3051)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGCTACATGGAGCGGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4624)	1		PPP4R1_+_9614519.23-P1P2	PPP4R1	P1P2	na	na	na	na	na	na	na	na	na
GAGCAGGGAGCAACGCACAC (SEQ ID NO: 1479)	TATGAACCACTAAGGCGTccaccttgttgGAGCAGGGAGCAACGCACACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3052)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTGTGCGTTGCTCCCTGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4625)	1		PRKDC_-_48872669.23-P1P2	PRKDC	P1P2	na	na	na	na	na	na	na	na	na

GGGCGGACTCGGCGGCA TGG (SEQ ID NO: 1480)	TATGAACCACTAAGGCGTcaacctgtggGGCGGGAGTCGGCGCGTATGG (SEQ ID NO: 3053)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCATGCCGCGCCATGTCCCGCCGCCAACAAA39TGGACGGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4626)	PRKDC_+_4 88726579 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCAGGGCAGCACGG CAG (SEQ ID NO: 1481)	TATGAACCACTAAGGCGTcaacctgtgGGGCCAGGGCAGCGTATGG (SEQ ID NO: 3054)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCTTGCGGTGCCTGGGCCGCCAACAAGGTTGACAGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4627)	PRKDC_-_48872617 2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTAGCGTTAGCCGTTGA CGT (SEQ ID NO: 1482)	TATGAACCACTAAGGCGTcaacctgtgGTTAGCGTTAGCGTATGG (SEQ ID NO: 3055)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCATGCCGTTAGTGGTTCATA (SEQ ID NO: 4628)	PRPF19_-_80674010 2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGCCGTCGGAGCACGG TGA (SEQ ID NO: 1483)	TATGAACCACTAAGGCGTcaacctgtgCGCGCCGTCGGAGCGTATGG (SEQ ID NO: 3056)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCATGCCGTTAGTGGTTCAT A (SEQ ID NO: 4629)	PRPF19_+_5 0673858 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGTTGCCGGGGACTGCT TCG (SEQ ID NO: 1484)	TATGAACCACTAAGGCGTcaacctgtgGTTGCCGGGGACGCGTATGG (SEQ ID NO: 3057)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCGGGAGCATCCGTCCGGGCAGCACCAGGGTCGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4630)	PRPF19_-_60679960 2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCACTAGCGGGAGAG AGG (SEQ ID NO: 1485)	TATGAACCACTAAGGCGTcaacctgtgGGGCACTAGAGCGCGTATGG (SEQ ID NO: 3058)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCCTCTCTCCGGCTAGTGGCCCAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4631)	TFPT_+_546 19065 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTAAGCCGCGAGGAG GAA (SEQ ID NO: 1486)	TATGAACCACTAAGGCGTcaacctgtgGGGTAAGCCGCGAGGAGTATGG (SEQ ID NO: 3059)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACTTCCTCCTCGGCGGCTTACCCCAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4632)	TFPT_+_546 19985 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAGCAGGATCGGTCCA CAG (SEQ ID NO: 1487)	TATGAACCACTAAGGCGTcaacctgtgGTAGCAGGATCGGTATGG (SEQ ID NO: 3060)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCCTGTGGACCGGATCCTGCTGCTACCGAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4633)	TFPT_-_546 18974 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGAGGGCGTCTCAAGG GCT (SEQ ID NO: 1488)	TATGAACCACTAAGGCGTcaacctgtgGCAGAGGGCGTCTATGG (SEQ ID NO: 3061)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACAGCCGTTGAGAGCGCCTCTGCCACAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4634)	PSIP1_-_155109923 2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCCGGGGACCAAGCGCA CTA (SEQ ID NO: 1489)	TATGAAGCACTAAGGCGTcaacctgtgCCGGGGACCAAGCGTATGG (SEQ ID NO: 3062)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACTAGCTGGGCTCCGCGGCCAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4635)	PSIP1_+_15 510952 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGCGCAGAGGCGTCT CAA (SEQ ID NO: 1490)	TATGAACCACTAAGGCGTcaacctgtgGTGGCGCAGAGGGCGTATGG (SEQ ID NO: 3063)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTTAAACTTGAGAGCGCCTCTGCCGCCAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4636)	PSIP1_-_155112918 2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACTCTGGCCCCTTGCAA GCAA (SEQ ID NO: 1491)	TATGAACCACTAAGGCGTcaacctgtgGAACCTCGCCCCTTATGG (SEQ ID NO: 3064)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACTTGTTGGAAAGGGGCGAAGGTCCAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4637)	PTN_+_1370 28220 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGTCATCGGTCCAAGA AGG (SEQ ID NO: 1492)	TATGAAGCACTAAGGCGTcaacctgtgCGAGTGATCGGTATGG (SEQ ID NO: 3065)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCCTTCTCGGACGGAGTGATCTCCGAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4638)	PTN_+_1370 253005 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTTTGCTACCGGCTCAG TCC (SEQ ID NO: 1493)	TATGAACCACTAAGGCGTcaacctgtgGGTTTGCTACCGGTATGG (SEQ ID NO: 3066)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACGGGAGCTCAGGGGTTAGCAACCGCCAAGCAACGGTTGACAGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4639)	PTN_-_137028280 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAGTGCCCGTGTCTCGAA GAG (SEQ ID NO: 1494)	TATGAACCACTAAGGCGTcaacctgtgAGGAGTGCCCGTGTATGG (SEQ ID NO: 3067)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACCTCTTCGAGACCGGCTCCGCCCAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4640)	RAD1_+_349 15822 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGCGGTCTCGCTCGTCT TCA (SEQ ID NO: 1495)	TATGAACCACTAAGGCGTcaacctgtgGAGGCTCGGTCGTATGG (SEQ ID NO: 3068)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACGTGAAGAGCCGAGCCGAGGGCGCCTCCAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4641)	RAD1_-_34915861 2 3-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGCCTCGGCTCGTCT TCA (SEQ ID NO: 1496)	TATGAACCACTAAGGCGTcaacctgtgGGCGCCTCGGCTATGG (SEQ ID NO: 3069)	TCACCAGCAGAAGTTTACGCAGGTTAGCTCTTAAACGTGAAGAGGAGCGAGGGCGCCTCCAACAAGGTGGACGCCTTAGTTGTCTGGTTGA (SEQ ID NO: 4642)	RAD1_+_349 15585 23-P1P2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCGAAGAGCGGAGGCGA CCA (SEQ ID NO: 1497)	TATGAAGCACTAAGGCGTcaacttgtgGGAAGAGCGGAG CTCCGCTCTTCGCCAACAAGGTGGACCGCTTAGTGGTCGTA (SEQ ID NO: 3070)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACTGGTCGC 65571 1 23-P1P2	RAD17_+_68	RAD17	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGACTAGCCTAGGTG AAA (SEQ ID NO: 1498)	TATGAAGCACTAAGGCGTcaacttgtgGGGGACTAGCCTA GGTGAAAgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3071)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACTTTCACCT AAGGTTAGTCCCCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4544)	RAD17_+_68	RAD17	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CGGGGAGCCCGGTACCTGC GAG (SEQ ID NO: 1499)	TATGAAGCACTAAGGCGTcaacttgtgGGGGAGGCCGTAC CTCCGAGgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3072)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACCTCCGGAG GTACCGGCCTCCCCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4545)	RAD17_-_ 68665557.2 3-P1P2	RAD17	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGGCCCCGAGGCTACT GGG (SEQ ID NO: 1500)	TATGAAGCACTAAGGCGTcaacttgtgCAGGCCGCGGAGCGC TAGTCCGGgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3073)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACCCCAGCTAG CCTCCGGCCGCTCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4546)	RAD18_+_90 26597.23-P1P2	RAD18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCGCTCCCGAGGATGC TGG (SEQ ID NO: 1501)	TATGAAGCACTAAGGCGTcaacttgtgGTCGCTCCCGAGGA TATCGTGGgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3074)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACCACGGATC CTCGGGACCCGACCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4547)	RAD18_-_ 9005092.23-P1P2	RAD18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGAGCGCCGGAGGCT AGT (SEQ ID NO: 1502)	TATGAAGCACTAAGGCGTcaacttgtgGTGGAGCGCCGTGA GGCTAGTgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3075)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACACTAGCCT CCGGCGCTCCACCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4548)	RAD18_+_90 05100.23-P1P2	RAD18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGGAAGCGCGCGGCTG TGG (SEQ ID NO: 1503)	TATGAAGCACTAAGGCGTcaacttgtgGAAGGAGCGCGCG GCTGTGGgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3076)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACCACCAGC CGCGCGCTCCTTCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4549)	RAD21_+_111 7986868.23-P1P2	RAD21	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAAGGAGTCCGCGCTCCC CGA (SEQ ID NO: 1504)	TATGAAGCACTAAGGCGTcaacttgtgGCAAGGAGTCCGCGC TCCCCGGAgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3077)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACTCGGGGGA GCCGGACTCCTCCCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4550)	RAD21_-_ 11178687.31-23-P1P2	RAD21	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGATTTGAACCGAGGA TTT (SEQ ID NO: 1505)	TATGAAGCACTAAGGCGTcaacttgtgGCGATTTGAACCGA GGTTCAAATCGCCGACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 3078)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACAAATCGCTC GGTTCAAATCGCCGACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4551)	RAD21_+_111 7897046.23-P1P2	RAD21	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCCTCGGCGCTAAGA TGG (SEQ ID NO: 1506)	TATGAAGCACTAAGGCGTcaacttgtgGCCGCCTCGGCGCTA AAGATGGgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3079)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACCCATCTTA GCGCCCAGGCGCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4552)	RAD23A_-_ 13056702.2 3-P1P2	RAD23A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGCGTGCCCCGGGAAGT GGT (SEQ ID NO: 1507)	TATGAAGCACTAAGGCGTcaacttgtgGTGGCGTGCCCCGG GAAGTGGTgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3080)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACGACCACTTC CGGGGGCAGGGCACCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4553)	RAD23A_-_ 13056866.2 3-P1P2	RAD23A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGCGGGGGCCCGAGC GACG (SEQ ID NO: 1508)	TATGAAGCACTAAGGCGTcaacttgtgGTGGCGGGGCCCG AAGCGACGgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3081)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACCGTCGGCT CGGCGCCCCGCCAACAAGGTGGACCGCTTAGTGGTTCAT A (SEQ ID NO: 4554)	RAD23A_+_ 13056741.23-P1P2	RAD23A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCTAGCTCACCACACA GGA (SEQ ID NO: 1509)	TATGAAGCACTAAGGCGTcaacttgtgCGCCGCTAGCTCAC CACAAGGAgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3082)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACTCCTGTTC TGTGAGCTAGCGGGGCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4555)	RAD23B_+_ 110045592 3-P1P2	RAD23B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAATCCCGCTAGCTCAC CAC (SEQ ID NO: 1510)	TATGAAGCACTAAGGCGTcaacttgtgGGAATCCGCTAGCTA CTCACCAGgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3083)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACGGTCGTGA GCTAGCGGATTCCCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4556)	RAD23B_+_ 110045584 2 3-P1P2	RAD23B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CGTTCTGCAGAAGGGCGCG AGG (SEQ ID NO: 1511)	TATGAAGCACTAAGGCGTcaacttgtgCGTTCTGCAGAAGG GCGCGAGGgttaagagcaaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3084)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACCCTCGCGCG CCTTCTGCAGACGCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4557)	RAD23B_+_ 110045623 3-P1P2	RAD23B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGGGGTTCCGCATTGTGG CTA (SEQ ID NO: 1512)	TATGAAGCACTAAGGCGTcaacttgtgGCGGGGTTCGCATT ATGGGACCCCGGCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 3085)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACTAGCCACCA RAD50_-_ 131992737 23-P1P2	RAD50_-_ 131992737 23-P1P2	RAD50	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGTTGCGGGGTTCGCA TTG (SEQ ID NO: 1513)	TATGAAGCACTAAGGCGTcaacttgtgGCGGTTGCGGGGTT ACCCCGGCAACCACGCGCCAACAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 3086)	TCACCAGCAGACTTTAGCGAGCTTAAGCTGTTCAAACGCAATGCG RAD50_-_ 131992731 23-P1P2	RAD50_-_ 131992731 23-P1P2	RAD50	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGGCGCCTGCGGATCA CGT (SEQ ID NO: 1514)	TATGAAGCACTAAG3GCTcaacctgtgGGCGCCTGCGT	TCAGCAGCAGAGTTTAGCGAGGTTAGGTCTTAAACAGGTGATC GGCAG3GCGCGCCAACAAGGTG3AGCGCCTAGTCGTTGATA (SEQ ID NO: 4660)	1	RAD56_+_13 1992655.23-P1P2	RAD50	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGCGCCGCACTCTCC TTA (SEQ ID NO: 1515)	TATGAACCACTAAGGCTcaacctgtgGAAGCGCCGCACT	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACTAAGGAG AGTGCGCGCCGTCGTTGCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4661)	1	RAD51_+_40 897418.23-P1P2	RAD51	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCAGCCCCTAAGGAGA GTG (SEQ ID NO: 1516)	TATGAACCACTAAGGCTcaacctgtgGCCCGAGCCCCTAAG	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACCAACTGTCC TTAGGGCTCGAGCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4662)	1	RAD51_-_ 40987434.2 3-P1P2	RAD51	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCAGAGACCGAGCCC CTA (SEQ ID NO: 1517)	TATGAACCACTAAGGCTcaacctgtgCGGCCAGAGAACCG	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACTAGGGCT CGGTCTCTCGGCCCAAGCAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4663)	1	RAD51_-_ 40987425.2 3-P1P2	RAD51	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACCTCACAGGCCGCA CCA (SEQ ID NO: 1518)	TATGAACCACTAAGGCTcaacctgtgCACCTCACAGGC GCGCACCAgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4664)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACTGGTGCG GCGTGTGAGGTGCCAAGAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4664)	1	RAD51AP1_+_ 46482103.2 3-P1P2	RAD51AP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGATACTTCTGAGCAACG GG (SEQ ID NO: 1519)	TATGAACCACTAAGGCTcaacctgtgGGATACTTCTGAGG AAGCGGgttaagagctaagctgtGGTAAAGTCTGCTGGTGA	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACCCCGTTG CTCAGAAGTATCCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4665)	1	RAD51AP1_+_ 46482266.2 3-P1P2	RAD51AP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGAGGGCCTCGGAA GTC (SEQ ID NO: 1520)	TATGAAGCACTAAGGCTcaacctgtgCGGCGGAGGGCTC CCAAAGTCgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4666)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAAGGACTTCG GACGCCTCCGCGCCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4666)	1	RAD51AP1_-_ 46466227.23-P1P2	RAD51AP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAACCCCACTTCCCCCA CAA (SEQ ID NO: 1521)	TATGAACCACTAAGGCTcaacctgtgGGAACCCCACTTCC CCACAAgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4667)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACTTGTGGG GAAGTTGGGGTTCCCCAAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4667)	1	RAD51B_+_ 88286566.23-P1P2	RAD51B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTGGGGAAAACTTGAA AGT (SEQ ID NO: 1522)	TATGAACCACTAAGGCTcaacctgtgGGGTGGGGAAAACT TGAAAGTgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4668)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACACTTTCAA GTTTCCCCACCCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4668)	1	RAD51B_-_ 88286566.23-P1P2	RAD51B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGTGTGGGGTTTCGATATT GG (SEQ ID NO: 1523)	TATGAAGCACTAAGGCTcaacctgtgGTGTGGGGTTTCG ATATTGgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4669)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACCCAATATC GAAACCCAGCGACCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4669)	1	RAD51B_-_ 88286667.2 3-P1P2	RAD51B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GACTGCTGAGGAACTCCTT	TATGAACCACTAAGGCTcaacctgtgGACTGCTGAGGAA AGCTCCTAGgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4670)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACCTAGGAG TTCCTCAAGCAGTCCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4670)	1	RAD51C_-_ 56777124.2 3-P1P2	RAD51C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGGGCTCAGCTGCTAA CCC (SEQ ID NO: 1525)	TATGAACCACTAAGGCTcaacctgtgGCAGGGCTCACCTG CTAACCCgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4671)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACGGGGTTAG CAAGGTGAGCCCTGCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4671)	1	RAD51C_+_ 56769390.23-P1P2	RAD51C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCGAGACGAACGAAGA CTG (SEQ ID NO: 1526)	TATGAACCACTAAGGCTcaacctgtgCGGCGAGACGAACG AAGACTGgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4672)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACCAGTCTTC GTTCTCTCGGCCCCAAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4672)	1	RAD51C_+_ 56777099.23-P1P2	RAD51C	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGGCGCGCGAATCC ACG (SEQ ID NO: 1527)	TATGAACCACTAAGGCTcaacctgtgGCTCGGCGCGCGAA TCCACGgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4673)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACGCGTGGA TCCGCGCGCGACCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4673)	1	RAD51D_+_ 33446221.23-P1P2	RAD51D	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGAGAGGAGGGCGGCA CCAA (SEQ ID NO: 1528)	TATGAACCACTAAGGCTcaacctgtgGAGAGGAGGGCG GCGCCAAgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4674)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACTTGGTGCC CGCCTCCGTCGTCGCAACAAGGTGGACGCCTAGTGGTTCATA (SEQ ID NO: 4674)	1	RAD51D_-_ 33446788.2 3-P1P2	RAD51D	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGGGTGGCACCAAGC GTA (SEQ ID NO: 1529)	TATGAACCACTAAGGCTcaacctgtgCGGCACCAAGCGG CGCACCGgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4675)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACTACCCTTG GTGCGGCCTCCCCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4675)	1	RAD51D_-_ 33446794.2 3-P1P2	RAD51D	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACACAGGGAGGTCTCGA TCT (SEQ ID NO: 1530)	TATGAACCACTAAGGCTcaacctgtgGCACACAGGGAGGC TCGATCTgttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 4676)	TCACCAGCAGAGACTTTAGCGAGGTTAGGTCTTAAACGAGATCGA GCTCGCCTGTGTGTGGCAACAAGGTGGAGCGCCTAGTGGTTCATA (SEQ ID NO: 4676)	1	RAD52_-_ 1068870.23-P1	RAD52	P1	na	na	na	na	na	na	na	na	na	na	na	na	na

GGCAGCGCGCGGTGCACACA (SEQ ID NO: 1531)	TATGAACCACTAAGGCGTccaccttgttgGGCAGCGCGCGGTGCACACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3104)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTGTGCACCGCGCGCTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4677)	1	RAD52_-_1058797.23-P1	RAD52	P1	na	na	na	na	na	na	na	na	na
GGCCCGAGGCGCGTAAGTGG (SEQ ID NO: 1532)	TATGAACCACTAAGGCGTccaccttgttgGGCCCGAGGCGCGTAAGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3105)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACTTACGCGCCTCGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4678)	1	RAD52_+_1058756.23-P1	RAD52	P1	na	na	na	na	na	na	na	na	na
GCACGCTCCCGGGAGGCTGA (SEQ ID NO: 1533)	TATGAACCACTAAGGCGTccaccttgttgGCACGCTCCCGGGAGGCTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3106)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCAGCCTCCCGGGAGCGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4679)	1	RAD52_-_1100281.23-P2	RAD52	P2	na	na	na	na	na	na	na	na	na
GGCCGCGTCACGCTCCCGGG (SEQ ID NO: 1534)	TATGAACCACTAAGGCGTccaccttgttgGgccgcgTCACGCTCCCGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3107)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGGAGCGTGACGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4680)	1	RAD52_-_1100274.23-P2	RAD52	P2	na	na	na	na	na	na	na	na	na
GGCCTCCCGGGAGCGTGACG (SEQ ID NO: 1535)	TATGAACCACTAAGGCGTccaccttgttgGGCCTCCCGGGAGCGTGAcggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3108)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTCACGCTCCCGGGAGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4681)	1	RAD52_+_1100254.23-P2	RAD52	P2	na	na	na	na	na	na	na	na	na
GCCGGCGGCGAGGGGATAGC (SEQ ID NO: 1536)	TATGAACCACTAAGGCGTccaccttgttgGCCGGCGGCGAGGGGATAGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3109)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCTATCCCCTCGCCGCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4682)	1	RAD54B_+_95487280.23-P1P2	RAD54B	P1P2	na	na	na	na	na	na	na	na	na
GGATCCTTGAGCTTCTCCGG (SEQ ID NO: 1537)	TATGAACCACTAAGGCGTccaccttgttgGGATCCTTGAGCTTCTCCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3110)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGAGAAGCTCAAGGATCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4683)	1	RAD54B_+_95487295.23-P1P2	RAD54B	P1P2	na	na	na	na	na	na	na	na	na
GGGGATAGCTGGTTACCAGA (SEQ ID NO: 1538)	TATGAACCACTAAGGCGTccaccttgttgGGGGATAGCTGGTTACCAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3111)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTGGTAACCAGCTATCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4684)	1	RAD54B_+_95487269.23-P1P2	RAD54B	P1P2	na	na	na	na	na	na	na	na	na
GGCGCGCTTTGGGAACAGGA (SEQ ID NO: 1539)	TATGAACCACTAAGGCGTccaccttgttgGGCGCGCTTTGGGAACAGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3112)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCTGTTCCCAAAGCGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4685)	1	RAD54L_-_46713461.23-P1P2	RAD54L	P1P2	na	na	na	na	na	na	na	na	na
GGCGGGAGGTTCGATTGACC (SEQ ID NO: 1540)	TATGAACCACTAAGGCGTccaccttgttgGGCGGGAGGTTCGATTGACCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3113)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGTCAATCGAACCTCCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4686)	1	RAD54L_-_46713404.23-P1P2	RAD54L	P1P2	na	na	na	na	na	na	na	na	na
GTAACGCGCGCTTTGGGAAC (SEQ ID NO: 1541)	TATGAACCACTAAGGCGTccaccttgttgGTAACGCGCGCTTTGGGAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3114)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTCCCAAAGCGCGCGTTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4687)	1	RAD54L_-_46713457.23-P1P2	RAD54L	P1P2	na	na	na	na	na	na	na	na	na
GAACAGGAGGAGTAAGATTG (SEQ ID NO: 1542)	TATGAACCACTAAGGCGTccaccttgttgGAACAGGAGGAGTAAGATTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3115)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAATCTTACTCCTCCTGTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4688)	1	RAD54L2_-_51575666.23-ENST000004	RAD54L2	ENST00000409535.2	na	na	na	na	na	na	na	na	na
GCTTACTCCTCCTGTTCCAA (SEQ ID NO: 1543)	TATGAACCACTAAGGCGTccaccttgttgGCTTACTCCTCCTGTTCCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3116)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGGAACAGGAGGAGTAAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4689)	1	RAD54L2_+_51575638.23-ENST00000409535.2	RAD54L2	ENST00000409535.2	na	na	na	na	na	na	na	na	na
GGAGTGCTGTCCCTTGGAAC (SEQ ID NO: 1544)	TATGAACCACTAAGGCGTccaccttgttgGGAGTGCTGTCCCTTGGAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3117)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTCCAAGGGACAGCACTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4690)	1	RAD54L2_-_51575650.23-ENST000004	RAD54L2	ENST00000409535.2	na	na	na	na	na	na	na	na	na
GACGGCCTTGCCGAGCACTG (SEQ ID NO: 1545)	TATGAACCACTAAGGCGTccaccttgttgGACGGCCTTGCCGAGCACTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3118)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGTGCTCGGCAAGGCCGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4691)	1	RAD9A_+_67159625.23-P1P2	RAD9A	P1P2	na	na	na	na	na	na	na	na	na
GAGAAGAGCGGAAGGGGGTT (SEQ ID NO: 1546)	TATGAACCACTAAGGCGTccaccttgttgGAGAAGAGCGGAAGGGGGTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3119)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAACCCCCTTCCGCTCTTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4692)	1	RAD9A_+_67159601.23-P1P2	RAD9A	P1P2	na	na	na	na	na	na	na	na	na
GGAAGGGACCCCGGACCCGG (SEQ ID NO: 1547)	TATGAACCACTAAGGCGTccaccttgttgGGAAGGGACCCCGGACCCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3120)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGGTCCGGGGTCCCTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4693)	1	RAD9A_-_67159466.23-P1P2	RAD9A	P1P2	na	na	na	na	na	na	na	na	na

GCCGCCATGGCCGAACAA GGA (SEQ ID NO: 1548)	TATGAAGCACTAAGGCGTcaacctgtgGCGGCCATGGCCG ACAAGCATggttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3121)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACTCCTTGTC GGCCATGGCGGCCGAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4694)	RBBP4_-_33116823.2 3-P1P2	RBBP7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGCGCAGGAAACAA TAG (SEQ ID NO: 1549)	TATGAACCACTAAGGCGTcaacctgtgGGGGGCGCAGGAA ACAATAGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3122)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACTCATTGTT TCCTCTGCGCCCCGCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4695)	RBBP4_-_33116821.2 3-P1P2	RBBP4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCGCAGTCCCCGGTAC TGA (SEQ ID NO: 1550)	TATGAACCACTAAGGCGTcaacctgtgGGTCGCAGTCCCC GGTACTGAgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3123)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACTCAGTACC GGGGACTGACCCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4696)	RBBP4_-_33117042.2 3-P1P2	RBBP4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTGTGACCGCGGTTGC GAA (SEQ ID NO: 1551)	TATGAACCACTAAGGCGTcaacctgtgCAGTGTGACCGCG TTGGGAAgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3124)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACTTCCCAAG CGGCGTCACACTCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4697)	RBBP7_+_16899257.23-P1P2	RBBP7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAACGAGAGGAGTCCTG AGG (SEQ ID NO: 1552)	TATGAACCACTAAGGCGTcaacctgtgGTAACGAGAGGAG TCCTGAGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3125)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCCTCGGA CTCCTCTCGTTACCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4698)	RBBP7_+_168992047.23-P1P2	RBBP7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGCGGGGTAGCGAGGT CTG (SEQ ID NO: 1553)	TATGAACCACTAAGGCGTcaacctgtgGTCGCGGGGTAGC GAGGTCTGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3126)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCAGACGTC GGTACCGCGCAGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4699)	RBBP7_-_18869863.23-P1P2	RBBP7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAATCCCGAGGCAATCTG ATC (SEQ ID NO: 1554)	TATGAAGCACTAAGGCGTcaacctgtgGAATCCCGAGGCA ATCTCGGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3127)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCCGAGATT CGCCTCGGGATTCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4700)	RBBP8_-_20513862.23-P1P2	RBBP8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCAGAACCCGGCAGGCG AAC (SEQ ID NO: 1555)	TATGAACCACTAAGGCGTcaacctgtgGCCAGAACCCGGCAC CGCGGAAgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3128)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACGTTCCGC GTCGCGGGTCTGTGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4701)	RBBP8_-_20513527.23-P1P2	RBBP8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGAGGCAATCTCGG AGG (SEQ ID NO: 1556)	TATGAACCACTAAGGCGTcaacctgtgGCCCGAGGCAATC TCGGAGGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3129)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCCTCCGA GATTGGCCTCGGGGCCAACAAGGTGCAGCTTAGTGGTTCATA (SEQ ID NO: 4702)	RBBP8_-_20513527.23-P1P2	RBBP8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGCAACGACGCCGCGAA CAA (SEQ ID NO: 1557)	TATGAAGCACTAAGGCGTcaacctgtgCGCGCGCAACGAA CACTCGTTgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3130)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACGTTGTTCGC CACTCGTTGCGCCGAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4703)	RBMX_-_136862895.23-P1P2	RBMX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTAGCACTACTGCGCAAC GA (SEQ ID NO: 1558)	TATGAACCACTAAGGCGTcaacctgtgGTAGCACTACTGC GCAAGTAGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3131)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACTGCTTGC GCAAGTAGTGCTACCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4704)	RBMX_-_136862895.23-P1P2	RBMX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTATATCTCCCTGGAGAT GG (SEQ ID NO: 1559)	TATGAACCACTAAGGCGTcaacctgtgGTATATCTCCCTGG AGATGGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3132)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCCATCTCC AGGGGAGATATAGCAAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4705)	RBMX_+_135928711.23-P1P2	RBMX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGTTAGCCGCGCAG GTG (SEQ ID NO: 1560)	TATGAAGCACTAAGGCGTcaacctgtgGAGGGTTAGCCGC GCAGGTGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3133)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCACCTGC GCGGCTAACCCCTCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4706)	RDM1_+_34257741.23-P1P2	RDM1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGTCCCCAGTCAGGG GGT (SEQ ID NO: 1561)	TATGAACCACTAAGGCGTcaacctgtgGGAAGTCCCCAGT CCGGGGTgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3134)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACGACCGGCT GACTGGGAGTTCCCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4707)	RDM1_-_34257593.23-P1P2	RDM1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTGCGCAGGCGCAATG CTG (SEQ ID NO: 1562)	TATGAACCACTAAGGCGTcaacctgtgGTTGCGCAGGCGC AATGCTGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3135)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACGAGCATTG CGCCTCGGGAACCAACAAGGTGCAGCTTAGTGGTTCATA (SEQ ID NO: 4708)	RDM1_+_34257785.23-P1P2	RDM1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGATCTAGTCCTGCAGTA GG (SEQ ID NO: 1563)	TATGAAGCACTAAGGCGTcaacctgtgGTGATCTAGTCCT GCAGTAGGgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3136)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCCTACTGC AGGATAGATCCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4709)	RECS_-_24641366.23-P1P2	RECS	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGTGCCCGTGGGCCAC CCG (SEQ ID NO: 1564)	TATGAACCACTAAGGCGTcaacctgtgGGCGTGCCCGTGG GCCACCCgttaagagctaagtgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3137)	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACGCGGTGTG GGCCCCGCGGGCCCGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4710)	RECS_-_24641324.23-P1P2	RECS	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na

GGGGCTTGCAGTGCGCTTCGG (SEQ ID NO: 1565)	TATGAAGCACTAAGGCGTcaacttgtgGGGCCTTGCACTGCAGG (SEQ ID NO: 3138)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCCCAAGC CACGTGCAACCGCCCCAACAAGCTGTGCTCGTTGATAGGTTCCATA (SEQ ID NO: 4711)	1	REG3_+_2463523-P1P2	REG3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGATCGGGAGAGTCGGGA (SEQ ID NO: 1566)	TATGAACCACTAAGGCGTcaacttgtgGAGATCGGGAGAGTCTTGCTGGTGA (SEQ ID NO: 3139)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACGGAGAGTTCCGG ACTGTCCGATCTGTCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4712)	1	REGQL_+_21654447 23-P1P2	REGQL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAGCCCGGTCTGTAAGTCTC (SEQ ID NO: 1567)	TATGAACCACTAAGGCGTcaacttgtgGAAGCCCGGTCTGTGGTGA (SEQ ID NO: 3140)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACGGAGAGTTCATA GAGCGGGTCTTCCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4713)	1	REGQL_-_21654423 23-P1P2	REGQL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTGTCGGAGATCCGAGAGT (SEQ ID NO: 1568)	TATGAACCACTAAGGCGTcaacttgtgCGTGTCGGAGATCCTGGTGA (SEQ ID NO: 3141)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCACTCTCGG GATGTCCGACCCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4714)	1	REGQL_+_21654469 23-P1P2	REGQL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGTCCGCGACGCGCGTT (SEQ ID NO: 1569)	TATGAACCACTAAGGCGTcaacttgtgGACGTCCGCGACGCTGGTGA (SEQ ID NO: 3142)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACTGGGACGCG GCTCGCGGGACGCTCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4715)	1	REGQL4_-_14674316923-P1P2	REGQL4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGCGCCGGAGGAGAGCT (SEQ ID NO: 1570)	TATGAACCACTAAGGCGTcaacttgtgGCGGCGCCGGAGGATGGTGA (SEQ ID NO: 3143)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCGGGGTCT CGTCCGGGCCGCCGCAACAAGGTTGGAGCGCCTTAGTTGGTTCAT A (SEQ ID NO: 4716)	1	REGQL4_+_14674296423-P1P2	REGQL4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCGAGAAGACGCGG (SEQ ID NO: 1571)	TATGAACCACTAAGGCGTcaacttgtgGggcgaGAGAGACGG CGTCTCGCGCCCCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4717)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACACCCGGT	1	REGQL4_+_14574294023-P1P2	REGQL4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTTTGGCTTCTCTCTGAGT (SEQ ID NO: 1572)	TATGAACCACTAAGGCGTcaacttgtgGCTTTGGCTTCTCTGGTTGA (SEQ ID NO: 3145)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCAACTCAG AGAAGCCAAACGCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4718)	1	REGQL5_+_73663200 23-P1P2	REGQL5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGTTGGGGTTGTGTA (SEQ ID NO: 1573)	TATGAACCACTAAGGCGTcaacttgtgGGAGTTGGGGTTGTGGTGA (SEQ ID NO: 3146)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCTATACAC AACCCCCAACTCCCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4719)	1	REGQL5_+_73663187 23-P1P2	REGQL5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACGTTGCGGAATTGAGC (SEQ ID NO: 1574)	TATGAAGCACTAAGGCGTcaacttgtgGGGCGTTGCGGAATGGTGA (SEQ ID NO: 3147)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACAAAGCTGA ATTCCCAAGCGCCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4720)	1	REGQL5_-_73663282 23-P1P2	REGQL5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCTCGCCAGGGGACCA (SEQ ID NO: 1575)	TATGAACCACTAAGGCGTcaacttgtgGCCCTCGCCAGGGGTGA AGCAGGGcgtaaagactaagctgCGTAAAGTCTCTGGTTGGTTGA (SEQ ID NO: 3148)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAAGCCCTGGT CCCTGGCGCAGGGGCCAACAAGGTTGGAGCGCCTTAGTTGGTTCAT A (SEQ ID NO: 4721)	1	REV1_-_100105381223-P1P2	REV1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGGGCGGGAGAGCC (SEQ ID NO: 1576)	TATGAACCACTAAGGCGTcaacttgtgGGAGGGGCGGGAGAGTGGTGA (SEQ ID NO: 3149)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCGGGTGGC TCTCGGCCCCTCCGCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4722)	1	REV1_+_100106487 23-P1P2	REV1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTGTGTTCCCGCTCAGCCG (SEQ ID NO: 1577)	TATGAAGCACTAAGGCGTcaacttgtgCGTTGGTTGCCTCCG AGCCGTGGgcttaagactaagctgCGTAAAGTCTGCTGGTTGA (SEQ ID NO: 3150)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAAGCCAGCGCT CGCGGGCAAACCAACCAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4723)	1	REV1_+_100108432 23-P1P2	REV1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCACGCGCCGGGAATCA (SEQ ID NO: 1578)	TATGAACCACTAAGGCGTcaacttgtgGCCACGCGCGGGAGTGGTGA (SEQ ID NO: 3151)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACGGTGTGATT CCCGCGCGGTGGCCAAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4724)	1	REV3L_-_111804372 23-P1P2	REV3L	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAGCCCTGTGAGCTTT (SEQ ID NO: 1579)	TATGAACCACTAAGGCGTcaacttgtgGGAAGCCCTGTGGTGA (SEQ ID NO: 3152)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACAGAAAG CTCGAGGGCGTTCCGCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4725)	1	REV3L_+_111804327 23-P1P2	REV3L	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGCCCTGTGATTCCC (SEQ ID NO: 1580)	TATGAAGCACTAAGGCGTcaacttgtgGTGAGCCCTGTGGTGA (SEQ ID NO: 3153)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCCGGGAA TTCCCCGGGttgtaagactaagctgCGTAAAGTCTCTGGTTGGTTCATA (SEQ ID NO: 4726)	1	REV3L_+_111804358 23-P1P2	REV3L	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACTTCGCGGTGAAGAAGCT (SEQ ID NO: 1581)	TATGAACCACTAAGGCGTcaacttgtgAAGTTCGCGGTGGTGGAG AAGTTCTGttgtaagactaagctgCGTAAAGTCTGCTGGTTGGTTCATA (SEQ ID NO: 3154)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACGACAACTTC TCCCGCGCGAAGTCCAACAAGGTTGGAGCGCCTTAGTTGGTTCATA (SEQ ID NO: 4727)	1	RFC1_+_38956734423-P1P2	RFC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGAATCTGTTATCGAGGC TC (SEQ ID NO: 1582)	TATGAAGCACTAAGGCGTcaacttgtggGAATCTGTTATCG	TCAGCAGCAGAGCTTTAGGCAGGTTAGGTCTTAAAGGAGCGTC GATGAGCACTTCCCAACAACGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4728)	1	RFC1_-_39367932.2 3-P1P2	RFC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCAAGAAGGCGAAGA CAG (SEQ ID NO: 1583)	TATGAACCACTAAGGCGTcaacttgtggTGCAAGAAGGCG	TCAACGAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACCTGTCTTC GGCTTCTTCGACCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4729)	1	RFC1_-_39367966.2 3-P1P2	RFC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CAAGAGGAGGCGGAACAA TGG (SEQ ID NO: 1584)	TATGAACCACTAAGGCGTcaacttgtgACGAGGAGGCGG	TCAACGAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACCCATCTC GCGGCTCCTCTCGAAGAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4730)	2	RFC2_+_736 69706.23-P1P2	RFC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGAAGGAGGCGCGA GAA (SEQ ID NO: 1585)	TATGAACCACTAAGGCGTcaacttgtgGGAAGAAGGAGG	TCAACGAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACTTCTCGCG GCTCTCTTCCCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4731)	2	RFC2_+_736 69709.23-P1P2	RFC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAGAGTGGAGAATGG AGG (SEQ ID NO: 1586)	TATGAACCACTAAGGCGTcaacttgtgGGAGAGTGGAGG	TCAACGAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACCCTCCATT CTCGGCGCTCCCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4732)	2	RFC2_+_736 69703.23-P1P2	RFC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGATACTTGTCCACCC AG (SEQ ID NO: 1587)	TATGAACCACTAAGGCGTcaacttgtgGCCGATACTTGTCC	TCAACGAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACCTGGGTG GACAAGTATCGGGCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4733)	2	RFC3_+_343 92320.23-P1P2	RFC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAACGGTCGGACTATC ACA (SEQ ID NO: 1588)	TATGAACCACTAAGGCGTcaacttgtgGGGACCGGTCGGAC	TCAACGAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACTGTGATAG TCCACGCCGTCCCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4734)	2	RFC3_-_343892375.2 3-P1P2	RFC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTGTCCACCCAGAAGGC TCA (SEQ ID NO: 1589)	TATGAACCACTAAGGCGTcaacttgtgGTTGTCCACCCAGA	TCACCAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACTGAGCCT CTTGGGTGAGACCAACCAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4735)	2	RFC3_+_343 892313.23-P1P2	RFC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACCTCCAGAAGGGGGCAT CAA (SEQ ID NO: 1590)	TATGAACCACTAAGGCGTcaacttgtgGACCTCCAGAAGGG	TCAACGAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACTTGATGCC GCTTGTGAGGTCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4736)	1	RFC4_-_196624629 23-P1P2	RFC4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCGAACTGAGGCGAC TGT (SEQ ID NO: 1591)	TATGAAGCACTAAGGCGTcaacttgtgCGCGGAACTGAGG	TCAGCAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACACAGTCG CCTCAGTTCCCGCGCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4737)	1	RFC4_+_1858 524228.23-P1P2	RFC4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTTCCGTGGCGGGAAA CTG (SEQ ID NO: 1592)	TATGAACCACTAAGGCGTcaacttgtgGCGTTCCGTGGCG	TCAGCAGCAGAGCTTTAGGCAGGTTAGGTCTTTAAACCAGTTCCC GCCAGCGGAAGCGCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4738)	1	RFC4_+_1858 524236.23-P1P2	RFC4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGTCTCAGGGTTCAGG TCG (SEQ ID NO: 1593)	TATGAACCACTAAGGCGTcaacttgtgGGGGTCTCAGGGT	TCAGCAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACCGGACCTG ACCCTGAGACCCCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4739)	1	RFC5_-_119454552 23-P1P2	RFC5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTCACTGTGCACGCG CTT (SEQ ID NO: 1594)	TATGAACCACTAAGGCGTcaacttgtgGGGTCACTGTGCA	TCAGCAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACAAGCGCC TGGACAGTGACCCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4740)	1	RFC5_-_119454586 23-P1P2	RFC5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTGACGCGACGATCT CAG (SEQ ID NO: 1595)	TATGAACCACTAAGGCGTcaacttgtgGGGTGACGCGACG	TCAGCAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACGCTGAGATC GTCCGCGTCACCCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4741)	1	RFC5_-_119454606 23-P1P2	RFC5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAGTTGGCGGTAGGTGC ATT (SEQ ID NO: 1596)	TATGAACCACTAAGGCGTcaacttgtgGCAGTTGGCGGTAG	TCAGCAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACAATGGAAGC TADGCCAACTGCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4742)	1	RFW01_+_74703751.23-P1P2	RFW01	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCTAGCGTTGCATTCGTGA ATTG (SEQ ID NO: 1597)	TATGAAGCACTAAGGCGTcaacttgtgGCGTAGGTTGCATT CGGGAGTGgttaagagctaagcgtgCGTAAAGTCTGCTGGTTGA (SEQ ID NO: 3170)	TCAGCAGCAGAGCTTTAGGCAGGTTAGGTCTTAAACGCACTCCGA ATGGACGTAGGGCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4743)	1	RFW03_+_74704744 23-P1P2	RFW03	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTTGAGCGCGCTCCGG AAG (SEQ ID NO: 1598)	TATGAACCACTAAGGCGTcaacttgtgGGCTTGAGCGCGCT	TCAACGAGCAGAGCTTTAGGCAGGTTAGGTCTTAAAGCTTCCGG AGGCGGTCAAGCCCCAACAAGGTTGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4744)	1	RFW03_+_74706773 23-P1P2	RFW03	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GACCGCAGGCCGGCTTGTGG (SEQ ID NO: 1599)	TATGAACCACTAAGGCGTccacctigttgGACCGCAGGCCGGCTTGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3172)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACAAGCCGGCCTGCGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4745)	1	RHNO1_-_2986478.23-P1P2	RHNO1	P1P2	TULP3_-_2986478.23-P1	TULP3	P1	na	na	na	na	na	na
GCAGGAGCTGCGGCCCCGTC (SEQ ID NO: 1600)	TATGAACCACTAAGGCGTccaccttgttgGCAGGAGCTGCGGCCCCGTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3173)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGACGGGGCCGCAGCTCCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4746)	1	RHNO1_-_2986438.23-P1P2	RHNO1	P1P2	TULP3_-_2986438.23-P1	TULP3	P1	na	na	na	na	na	na
GCGTCCGGCCCGGGCTGGTA (SEQ ID NO: 1601)	TATGAACCACTAAGGCGTccaccttgttgGCGTCCGGCCCGGGCTGGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3174)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACCAGCCCGGGCCGGACGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4747)	1	RHNO1_-_2986453.23-P1P2	RHNO1	P1P2	TULP3_-_2986453.23-P1	TULP3	P1	na	na	na	na	na	na
GAAGCTGCCCGGGAAAGTGG (SEQ ID NO: 1602)	TATGAACCACTAAGGCGTccaccttgttgGAAGCTGCCCGGGAAAGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3175)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACTTTCCCGGGCAGCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4748)	1	RIF1_-_15226670023-P1P2	RIF1	P1P2	na	na	na	na	na	na	na	na	na
GGAGGGAGGAGATCGGCGGA (SEQ ID NO: 1603)	TATGAACCACTAAGGCGTccaccttgttgGGAGGGAGGAGATCGGCGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3176)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCGCCGATCTCCTCCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4749)	1	RIF1_+_152266651.23-P1P2	RIF1	P1P2	na	na	na	na	na	na	na	na	na
GGCGAGATGCCTCGTTCCCC (SEQ ID NO: 1604)	TATGAACCACTAAGGCGTccaccttgttgGGCGAGATGCCTCGTTCCCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3177)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGGAACGAGGCATCTCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4750)	1	RIF1_-_152266641.23-P1P2	RIF1	P1P2	na	na	na	na	na	na	na	na	na
GGAGAAGGCGAAGAAGGCTG (SEQ ID NO: 1605)	TATGAACCACTAAGGCGTccaccttgttgGGAGAAGGCGAAGAAGGCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3178)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGCCTTCTTCGCCTTCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4751)	1	RING1_+_33176535.23-P1P2	RING1	P1P2	na	na	na	na	na	na	na	na	na
GGCCGAGGAGAAGGCGAAGA (SEQ ID NO: 1606)	TATGAACCACTAAGGCGTccaccttgttgGGCCGAGGAGAAGGCGAAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3179)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTTCGCCTTCTCCTCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4752)	1	RING1_+_33176541.23-P1P2	RING1	P1P2	na	na	na	na	na	na	na	na	na
GGCTGTCTGAGCAGCGGTTG (SEQ ID NO: 1607)	TATGAACCACTAAGGCGTccaccttgttgGGCTGTCTGAGCAGCGGTTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3180)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAACCGCTGCTCAGACAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4753)	1	RING1_-_33176343.23-P1P2	RING1	P1P2	na	na	na	na	na	na	na	na	na
GGAAGTGCGGTGTTGAGCGC (SEQ ID NO: 1608)	TATGAACCACTAAGGCGTccaccttgttgGGAAGTGCGGTGTTGAGCGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3181)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGCTCAACACCGCACTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4754)	1	RNASEH1_+_3605908.23-P1P2	RNASEH1	P1P2	na	na	na	na	na	na	na	na	na
GGCGGCCAAGGCGACTCTGT (SEQ ID NO: 1609)	TATGAACCACTAAGGCGTccaccttgttgGGCGGCCAAGGCGACTCTGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3182)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACAGAGTCGCCTTGGCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4755)	1	RNASEH1_-_3605827.23-P1P2	RNASEH1	P1P2	na	na	na	na	na	na	na	na	na
GGGAGAGAAGGGGCCCAACA (SEQ ID NO: 1610)	TATGAACCACTAAGGCGTccaccttgttgGGGAGAGAAGGGGCCCAACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3183)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTTGGGCCCCTTCTCTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4756)	1	RNASEH1_+_3606182.23-P1P2	RNASEH1	P1P2	na	na	na	na	na	na	na	na	na
GAGCTGGAGAGAGACAATAC (SEQ ID NO: 1611)	TATGAACCACTAAGGCGTccaccttgttgGAGCTGGAGAGAGACAATACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3184)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTATTGTCTCTCTCCAGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4757)	1	RNASEH2A_-_12917521.23-P1P2	RNASEH2A	P1P2	na	na	na	na	na	na	na	na	na
GAGGCGAACTCAGGCGACAG (SEQ ID NO: 1612)	TATGAACCACTAAGGCGTccaccttgttgGAGGCGAACTCAGGCGACAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3185)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGTCGCCTGAGTTCGCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4758)	1	RNASEH2A_+_12917522.23-P1P2	RNASEH2A	P1P2	na	na	na	na	na	na	na	na	na
GCTCCCCTCCCTGGCCCTAG (SEQ ID NO: 1613)	TATGAACCACTAAGGCGTccaccttgttgGCTCCCCTCCCTGGCCCTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3186)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTAGGGCCAGGGAGGGGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4759)	1	RNASEH2A_+_12917619.23-P1P2	RNASEH2A	P1P2	na	na	na	na	na	na	na	na	na
GCTGGCTGCGTAGGAGGGAC (SEQ ID NO: 1614)	TATGAACCACTAAGGCGTccaccttgttgGCTGGCTGCGTAGGAGGGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3187)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCCCTCCTACGCAGCCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4760)	1	RNF168_+_196230575.23-P1P2	RNF168	P1P2	na	na	na	na	na	na	na	na	na
GGCGCCCGAAAGGTGAACAC (SEQ ID NO: 1615)	TATGAACCACTAAGGCGTccaccttgttgGGCGCCCGAAAGGTGAACACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3188)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTGTTCACCTTTCGGGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4761)	1	RNF168_+_196230525.23-P1P2	RNF168	P1P2	na	na	na	na	na	na	na	na	na

GTGCGCTGCTGCCGAAGGA ATA (SEQ ID NO: 1616)	TATGAACCACTAAGGCGTcaacctgtgGTGCGCTGCCGGA ACCGAATAgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3189)	TCACCAGCAGAGACTTTAGCGCAGCTTAAACTATTGGTT GCCGGCAGCGGACCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4762)	1: 95230800 23-P1P2	RNF168_+_1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCGCGCGCCGGCAGA ATA (SEQ ID NO: 1617)	TATGAACCACTAAGGCGTcaacctgtgGCCGCGCGCCGG CACAATAgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3190)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTATTGTGC GGCGTGCGCGCGGCAAGAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4763)	1: 74687 23-P1P2	RNF2_+_185	RNF2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGCGCCGGCAGGAGA CTC (SEQ ID NO: 1618)	TATGAACCACTAAGGCGTcaacctgtgGCGCGCGCCGGA GAAGACTCgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3191)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACGAAGTCTTG CTCGCGCGCGCGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4764)	1: 914727 23-P1P2	RNF2_+_185	RNF2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGCGCTCCGCCAATATT GTG (SEQ ID NO: 1619)	TATGAACCACTAAGGCGTcaacctgtgGTGAGCGCTCCGCCA TATTGTCgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3192)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACCACAATAT GGCGAGCGCTCTCACCAACAAGGTGCAGCTTAGTGGTTCATA (SEQ ID NO: 4765)	1: 98614639 23-P1P2	RNF2_-	RNF2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCTGGCAGAGTCCT GTT (SEQ ID NO: 1620)	TATGAACCACTAAGGCGTcaacctgtgGGGCCTGGCAGAG TCCTGTTgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3193)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACAAGGA CTTGTGCCAGGCGCCCAAGAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4766)	1: 42962693 23-P1P2	RNF20_+_103	RNF20	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCGAGTTCTGTTCCCC GCG (SEQ ID NO: 1621)	TATGAACCACTAAGGCGTcaacctgtgGTGCGAGTTCTGTT CCCGCGgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3194)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACCCGCGGGA ACAGAACTCGGCAGCAAACAAGGTGCAGCTTAGTGGTTCATA (SEQ ID NO: 4767)	1: 104296229 23-P1P2	RNF20_-	RNF20	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGTCTCGCGCCGCGGGT CAG (SEQ ID NO: 1622)	TATGAACCACTAAGGCGTcaacctgtgGTGTCTCGCGCCGC GGGTCAGgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3195)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACGTCGGCG GGCGCGCGAGACACCAAAGGTGCAGCTCTTAGTGGTTCAT A (SEQ ID NO: 4768)	1: 104296166 23-P1P2	RNF20_-	RNF20	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGCGCGGCATCTTTC TCG (SEQ ID NO: 1623)	TATGAACCACTAAGGCGTcaacctgtgGGGCGCGCGGCATC TTTCTCGgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3196)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACCGAGAAA GATGCGCGCGCGGCCCAAGAAGGTGCAGCTCTTAGTGGTTCAT A (SEQ ID NO: 4769)	1: 24708175 23-P1P2	RNF4_-	RNF4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTAGGAGGGAAGAGCGCG GGC (SEQ ID NO: 1624)	TATGAACCACTAAGGCGTcaacctgtgGTTAGGAGGGAAGA GCGCGGgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3197)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACGCGCCGGG TCGTTCCCGCCTAACCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4770)	1: 24713492 23-P1P2	RNF4_+_	RNF4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTTCTCGAAGCAGCTCTC CT (SEQ ID NO: 1625)	TATGAACCACTAAGGCGTcaacctgtgGTTTCTCGAAGCAG CTCTCCTgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3198)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACAGGAGAAG CTCCTCGGAGAAACCAACAAGGTGCAGCTTAGTGGTTCATA (SEQ ID NO: 4771)	1: 24708837 23-P1P2	RNF4_-	RNF4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGGTGAAATCCAAGATT GG (SEQ ID NO: 1626)	TATGAACCACTAAGGCGTcaacctgtgGCCGGGTGAAATCC AAGATTGgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3199)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACCCATCTTG GATTTCACCGGGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4772)	1: 30773556 2 3-P1P2	RNF40_-	RNF40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTCCGGTGAAATCCAA AGA (SEQ ID NO: 1627)	TATGAACCACTAAGGCGTcaacctgtgGCGTCCGGTGAAA TCCAAGAgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3200)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACTCTTGGAT TTCAACGGACGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4773)	1: 30773555 2 3-P1P2	RNF40_-	RNF40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCTACGCGTCGACCTCC TGC (SEQ ID NO: 1628)	TATGAACGCACTAAGGCGTcaacctgtgCGCTAGGGTCGACC CTCCTGGgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3201)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACGGGACAG GGTCGACGCCTAGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4774)	1: 30773861 2 3-P1P2	RNF40_-	RNF40	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCGCGCGGTGGCCTG CCG (SEQ ID NO: 1629)	TATGAACCACTAAGGCGTcaacctgtgGCCGCGCGCGGTGG CCTGCCGgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3202)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACGGCGTCCTG CCACCAGCTCGCGGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4775)	1: 37321993 2 3-P1P2	RNF8_-	RNF8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGGCAGCTGGTGCCTG CGG (SEQ ID NO: 1630)	TATGAACCACTAAGGCGTcaacctgtgAGCGGCAGCTGGTG CCTGCGGgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3203)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACGCAGGG CACCAGCTCGCGGCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4776)	1: 37322035 2 3-P1P2	RNF8_-	RNF8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCAGCCAGAAGCTAGG TCA (SEQ ID NO: 1631)	TATGAACCACTAAGGCGTcaacctgtgGTGCAGCCAGAAGC TAGGTCAgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3204)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACGTCGACCTAG GTTCTGGCTGCCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4777)	1: 37321950 2 3-P1P2	RNF8_-	RNF8	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGAGCGGCGGATGAACC GCC (SEQ ID NO: 1632)	TATGAACCACTAAGGCGTcaacctgtgGCGAGCGGCGGAT GAACCGCgttaagagctaagtgtGTAAAGTCTGCTGGTGA (SEQ ID NO: 3205)	TCACCAGCAGAGACTTTAGCGCAGCTTAAGCTGTTAAACGGCGGTTC ATCCGCCGTCCGCCCAACAAGGTGCAGCTCTTAGTGGTTCATA (SEQ ID NO: 4778)	1: 1733528 23-P1P2	RPA1_-	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGGCGGTGCGGCGAACTTCTC (SEQ ID NO: 1633)	TATGAAGCACTAAGGCGTcaacttgtgGGCGGTGCGCGAGTTTGTCGTgattaagagtaagtgtCGTAAAGTCTGCTCGTTGA (SEQ ID NO: 3296)	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGAAGAGTTT	1	1733203.23-P1P2	RPA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGCTGGAGCTGTTGCG (SEQ ID NO: 1634)	TATGAACCACTAAGGCGTcaacttgtgGGAAAGCTGGAGCTGTTGA (SEQ ID NO: 3297)	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGGCAACA	1	RPA1_-_1733388.23-P1P2	RPA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGCAAGAGCGGCGCAGTTGG (SEQ ID NO: 1635)	TATGAACCACTAAGGCGTcaacttgtgAAGCGCACAAGCGCC	TCAACAGCAGAGCTTACGCAGCTTAGCTCTTAAACCCACTGC	2	RPA2_+_282.41028.23-P1P2	RPA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTAGCGACGAGCGCGCCAG (SEQ ID NO: 1636)	TATGAACCACTAAGGCGTcaacttgtgGAGTAGCGACGAGCC	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACCTGCGCC	2	RPA2_+_282.41031.23-P1P2	RPA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCTCTGGCTAGTTTCTTC (SEQ ID NO: 1637)	TATGAACCACTAAGGCGTcaacttgtgGCGCTCTGGCTAC	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGAGAAA	2	RPA2_-_29241061.23-P1P2	RPA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAATCGCTGTGATCGCA (SEQ ID NO: 1638)	TATGAACCACTAAGGCGTcaacttgtgGGAATCGCTGTGATC	TCAACAGCAGAGCTTACGCAGCTTAGCTCTTAAACTTCTGCGA	2	RPA3_+_7381.6362.23-P1P2	RPA3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGATCACAGGAATTCCC (SEQ ID NO: 1639)	TATGAACCACTAAGGCGTcaacttgtgCGGCGAATCACAGGA	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACCCGGGAA	2	RPA3_-_7380380.23-P1P2	RPA3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGCGCTGGCGGGAAACCA (SEQ ID NO: 1640)	TATGAACCACTAAGGCGTcaacttgtgGGCGGCTGGCGGGAG	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACTGGGTTTC	2	RPA3_-_7680087.23-P1P2	RPA3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGCTTCGAGCTGTGCC (SEQ ID NO: 1641)	TATGAACCACTAAGGCGTcaacttgtgGAGGGCTTCGAGCT	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACTTCGGACA	1	RPA4_+_961.39196.23-P1P2	RPA4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCAGGGCTGAGGAGAA (SEQ ID NO: 1642)	TATGAAGCACTAAGGCGTcaacttgtgGCTCAGGGCTGAGGA	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACAGCCTCTC	1	RPA4_+_961.39276.23-P1P2	RPA4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGTCTGGGGCTCCAC (SEQ ID NO: 1643)	TATGAACCACTAAGGCGTcaacttgtgGTCGTCTGGGGCT	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCCGTGG	1	RPA4_+_961.39247.23-P1P2	RPA4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAGCGGGCCGGGCTG (SEQ ID NO: 1644)	TATGAACCACTAAGGCGTcaacttgtgGGAAGCGGGCCGGGG	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGGATCCAG	1	RRM2B_+_1.63251059.23-P1P2	RRM2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAGAAACCCTTAGGC (SEQ ID NO: 1645)	TATGAAGCACTAAGGCGTcaacttgtgCGggaaaaccctaggcc	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGCCGCCT	1	RRM2B_+_1.63251857.23-P1P2	RRM2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCGTGTCGTAGTTGG (SEQ ID NO: 1646)	TATGAACCACTAAGGCGTcaacttgtgGGTCGTCTCTAC	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGCCTCGAAC	1	RRM2B_+_1.63251153.23-P1P2	RRM2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGGGTTGGACGGCCCTC (SEQ ID NO: 1647)	TATGAACCACTAAGGCGTcaacttgtgGACGGGTTGGCGTGA	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGTTCGAGG	1	RTEL1_-_82289757.23-P1P2	RTEL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGGCTGACAGGTTGGG (SEQ ID NO: 1648)	TATGAAGCACTAAGGCGTcaacttgtgCGTGGCTGACAGCC	TCAGCAGCAGAGCTTACGCAGCTTAGCTCTTAAACGTCCCCA	1	RTEL1_-_82289740.23-P1P2	RTEL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCAGGTCGAGAGTTTTCC (SEQ ID NO: 1649)	TATGAACCACTAAGGCGTcaacttgtgGCCAGGTCGAGAG	CTCTCTGACGCTCGGCCAACAAGGTGGAACGGCTAGTGCTTCATA	1	RTEL1_+_82289702.23-P1P2	RTEL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GAACGGCCGCATGGTAACTC (SEQ ID NO: 1650)	TATGAACCACTAAGGCGTccaccttgttgGAACGGCCGCATGGTAACTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3223)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGTTACCATGCGGCCGTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4796)	2	RUVBL1_+_127842630.23-P1P2	RUVBL1	P1P2	na	na	na	na	na	na	na	na	na
GCGCATGGTAACTCAGGCGC (SEQ ID NO: 1651)	TATGAACCACTAAGGCGTccaccttgttgGCGCATGGTAACTCAGGCGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3224)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGCCTGAGTTACCATGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4797)	2	RUVBL1_+_127842624.23-P1P2	RUVBL1	P1P2	na	na	na	na	na	na	na	na	na
GGGCCAGGAGAACGCGCGAG (SEQ ID NO: 1652)	TATGAACCACTAAGGCGTccaccttgttgGGGCCAGGAGAACGCGCGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3225)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCGCGCGTTCTCCTGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4798)	2	RUVBL1_+_127842425.23-P1P2	RUVBL1	P1P2	na	na	na	na	na	na	na	na	na
GCATGCACCCACACTTACAA (SEQ ID NO: 1653)	TATGAACCACTAAGGCGTccaccttgttgGCATGCACCCACACTTACAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3226)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGTAAGTGTGGGTGCATGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4799)	2	RUVBL2_+_49497175.23-P1P2	RUVBL2	P1P2	na	na	na	na	na	na	na	na	na
GTGCGCTTTACGGAAACGAG (SEQ ID NO: 1654)	TATGAACCACTAAGGCGTccaccttgttgGTGCGCTTTACGGAAACGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3227)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCGTTTCCGTAAAGCGCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4800)	2	RUVBL2_-_49496777.23-P1P2	RUVBL2	P1P2	na	na	na	na	na	na	na	na	na
GTGGCAGTTGGTGAGCATCA (SEQ ID NO: 1655)	TATGAACCACTAAGGCGTccaccttgttgGTGGCAGTTGGTGAGCATCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3228)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGATGCTCACCAACTGCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4801)	2	RUVBL2_-_49497171.23-P1P2	RUVBL2	P1P2	na	na	na	na	na	na	na	na	na
GCGGTTGGCTTGAGCGGGAC (SEQ ID NO: 1656)	TATGAACCACTAAGGCGTccaccttgttgGCGGTTGGCTTGAGCGGGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3229)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCCCGCTCAAGCCAACCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4802)	2	SAE1_-_47634161.23-P1P2	SAE1	P1P2	na	na	na	na	na	na	na	na	na
GGGAGGCGGCACAGTATGAC (SEQ ID NO: 1657)	TATGAACCACTAAGGCGTccaccttgttgGGGAGGCGGCACAGTATGACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3230)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTCATACTGTGCCGCCTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4803)	2	SAE1_-_47634249.23-P1P2	SAE1	P1P2	na	na	na	na	na	na	na	na	na
GTTGAGCGGGACCGGAGCTG (SEQ ID NO: 1658)	TATGAACCACTAAGGCGTccaccttgttgGTTGAGCGGGACCGGAGCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3231)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGCTCCGGTCCCGCTCAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4804)	2	SAE1_-_47634169.23-P1P2	SAE1	P1P2	na	na	na	na	na	na	na	na	na
GAAGGGGAATTGGTAAACCA (SEQ ID NO: 1659)	TATGAACCACTAAGGCGTccaccttgttgGAAGGGGAATTGGTAAACCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3232)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGTTTACCAATTCCCCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4805)	1	SCMH1_+_41626493.23-ENST00000326197.7	SCMH1	ENST00000326197.7	na	na	na	na	na	na	na	na	na
GCTGCTGACATCCTGGATCT (SEQ ID NO: 1660)	TATGAACCACTAAGGCGTccaccttgttgGCTGCTGACATCCTGGATCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3233)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGATCCAGGATGTCAGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4806)	1	SCMH1_+_41626545.23-ENST00000326197.7	SCMH1	ENST00000326197.7	na	na	na	na	na	na	na	na	na
GTGAAGTACATGGGATAATT (SEQ ID NO: 1661)	TATGAACCACTAAGGCGTccaccttgttgGTGAAGTACATGGGATAATTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3234)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAATTATCCCATGTACTTCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4807)	1	SCMH1_-_41626405.23-ENST000003	SCMH1	ENST00000326197.7	na	na	na	na	na	na	na	na	na
GAGTCTGCATCCCAATATCA (SEQ ID NO: 1662)	TATGAACCACTAAGGCGTccaccttgttgGAGTCTGCATCCCAATATCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3235)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGATATTGGGATGCAGACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4808)	1	SCMH1_+_41627034.23-ENST00000402904.2	SCMH1	ENST00000402904.2	na	na	na	na	na	na	na	na	na
GCTCCAGTCTATCACTGCAG (SEQ ID NO: 1663)	TATGAACCACTAAGGCGTccaccttgttgGCTCCAGTCTATCACTGCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3236)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGCAGTGATAGACTGGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4809)	1	SCMH1_-_41627112.23-ENST000004	SCMH1	ENST00000402904.2	na	na	na	na	na	na	na	na	na
GGCATCCCAATATCAAGGTA (SEQ ID NO: 1664)	TATGAACCACTAAGGCGTccaccttgttgGGCATCCCAATATCAAGGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3237)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACCTTGATATTGGGATGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4810)	1	SCMH1_+_41627029.23-ENST00000402904.2	SCMH1	ENST00000402904.2	na	na	na	na	na	na	na	na	na
GCAGCCGCCTCCGAAGATGG (SEQ ID NO: 1665)	TATGAACCACTAAGGCGTccaccttgttgGCAGCCGCCTCCGAAGATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3238)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATCTTCGGAGGCGGCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4811)	1	SETD2_+_47205374.23-P1P2	SETD2	P1P2	na	na	na	na	na	na	na	na	na
GGAGGCCGCAGGTCCGACCG (SEQ ID NO: 1666)	TATGAACCACTAAGGCGTccaccttgttgGGAGGCCGCAGGTCCGACCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3239)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGTCGGACCTGCGGCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4812)	1	SETD2_-_47205505.23-P1P2	SETD2	P1P2	na	na	na	na	na	na	na	na	na

G3GCAGGTGCTTCATTCGG CAA3 (SEQ ID NO: 1667)	TATGAAGCACTAAG3GCGTcaacctgtgG3GCAGCTGCTTCAT GAAG3CAGCTGCCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 3240)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCTCCGCAT 1 GAAG3CAGCTGCCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4813)	SEIT2_-47205422 23-P1P2	SEIT2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACATCCAGCTGCTCAGT CG (SEQ ID NO: 1668)	TATGAACCACTAAGGCGTcaacctgtgGACATCCAGCTGCT GCATGTGgttaagagctaagctgGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3241)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCTCCACAA 1 GCATGTGGTCGTCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4814)	SEITMAR_+_4346121 23-P1P2	SEITMAR	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGACACGGCGTTTGTG GGA (SEQ ID NO: 1669)	TATGAACCACTAAGCGCGTcaacctgtgGACGACACGGCGT TTGTGGGAgttaagagctaagctgGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3242)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCTCCACAA 1 GGCGGTGTCGTCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4815)	SEITMAR_-4346596 23-P1P2	SEITMAR	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTTCAGGAACACCACG TGG (SEQ ID NO: 1670)	TATGAACCACTAAGGCGTcaacctgtgCGCTTCAGCAGACC CACGTGCggttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3243)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCCACCGTG 1 GTCTCCTCAAGCCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4816)	SEITMAR_+_4346252 23-P1P2	SEITMAR	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGAAGCGGGAGCCGGTA TGG (SEQ ID NO: 1671)	TATGAACCACTAAGGCGTcaacctgtgGAGAGCGGGAGCCC CGTATGGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3244)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCATCCGG 1 GGCCCGCGCTCTCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4817)	SEITX_+_135123031 23-P1P2	SEITX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTAGAGCGGGGGCCCG GTA (SEQ ID NO: 1672)	TATGAACCACTAAGGCGTcaacctgtgGCCTAGAGCGGGGG CCCGGTAgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3245)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTACCGGG 1 CCCCGCTCTCTAGGCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4818)	SEITX_+_135123036 23-P1P2	SEITX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGACTGCCACGCCACA GAG (SEQ ID NO: 1673)	TATGAACCACTAAGGCGTcaacctgtgGGCGACGCCACAGC CACAGAGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3246)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCTCTCGG 1 CTGGCGCTCGGCCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4819)	SEITX_+_135123036 23-P1P2	SEITX	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGCGGTAGTTCCAGAG CGA (SEQ ID NO: 1674)	TATGAACCACTAAGGCGTcaacctgtgGCGGCGGTAGTTC CAGAGCGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3247)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCGGCCTG 2 GAACTACCGCGCGGCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4820)	SEGOL1_+_20227648 23-P1P2	SEGOL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGTAGTTCCAGGCGA CGG (SEQ ID NO: 1675)	TATGAACCACTAAGGCGTcaacctgtgGCGGTAGTTCCAGG CGGAACGGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3248)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCGTGCC 2 CTGGGAACTACCGCGCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4821)	SEGOL1_+_20227665 23-P1P2	SEGOL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCAGCAGGGCGGACCGGT GGTA (SEQ ID NO: 1676)	TATGAAGCACTAAGGCGTcaacctgtgGGCAGCAGGGCGG ACCGGTGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3249)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTACCACCTG 2 TCCGCCGTCGCCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4822)	SEGOL1_+_20227653 23-P1P2	SEGOL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGAATCCCGGAAGGAC AGA (SEQ ID NO: 1677)	TATGAACCACTAAGGCGTcaacctgtgGCCGGAATCCCGGA CGACAGAgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3250)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCTGTCCT 1 CCGGGGATCCGGCCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4823)	SEGOL2_+_20191389 23-P1P2	SEGOL2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCATGCGGCGCAACCGC CCA (SEQ ID NO: 1678)	TATGAACCACTAAGGCGTcaacctgtgGCTCATGCGGCGC CCCGCCAAgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3251)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTTGGCGGG 1 TTGCCGCATCAGCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4824)	SEGOL2_-20139098923-P1P2	SEGOL2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCCAACCGCCCAACCGG CTG (SEQ ID NO: 1679)	TATGAAGCACTAAGGCGTcaacctgtgCGGCCAACCGCCCA ACGGCTGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3252)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCGCCT 1 TGGGGGTTTGCCCCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4825)	SEGOL2_-20139098923-P1P2	SEGOL2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTCAGGGCTCCGTTG TGG (SEQ ID NO: 1680)	TATGAACCACTAAGGCGTcaacctgtgGGACTCAGGGCTCC CGTTGTGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3253)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGCCAGCACG 2 GAAGCCGTGAGTCGCCAAGAAG3GTGGACGCCGCTTACTCGTTCATA (SEQ ID NO: 4826)	SHFM1_+_9838933 23-P1P2	SHFM1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGTCGGAGGCGGGTAG TGA (SEQ ID NO: 1681)	TATGAACCACTAAGGCGTcaacctgtgAGGCGTCGGGAGGG CGTAGTGAgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3254)	TCAACAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCAGTACC 2 GCCTCCGGACCGCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4827)	SHFM1_+_96339150 23-P1P2	SHFM1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGGAGCGCGGTAAGTGA CCG (SEQ ID NO: 1682)	TATGAAGCACTAAGGCGTcaacctgtgGTCGGAGCGCGGT GTGACGGGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3255)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCGTCACT 2 ACCGCCTCCGACCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4828)	SHFM1_+_96339147 23-P1P2	SHFM1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCCGGCAAGCTAGGCG CTG (SEQ ID NO: 1683)	TATGAACCACTAAGGCGTcaacctgtgGGCCCGGCAAGCTA CGCTGCTGgttaagagctaagctgCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3256)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGACGGCGC 1 ACGCTGCCGGGGGCCCAACAAG3GTGGACGGCTTAGTCGTTCATA (SEQ ID NO: 4829)	SHFRH_+_146238458 23-P1P2	SHFRH	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTGGAGCCCGGGAAGCT CAA (SEQ ID NO: 1684)	TATGAACCACTAAGGCGTcaacttgtgGTGGAGCCCGGGA (SEQ ID NO: 3257)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTTGAGCTTACCCGGGCTCCACCAAGAAGGTTGGACCGCCTTACTTGTTCATA	SHPRH_+_46285768 23-P1P2	SHPRH	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGTTTCTGGGTCTCCC GG (SEQ ID NO: 1685)	TATGAACCACTAAGGCGTcaacttgtgTTGGTTTCTGTGGGTC (SEQ ID NO: 3258)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCGGGAG ACCCAGAGAACCACCAAGCAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4331)	SHPRH_+_46286438 23-P1P2	SHPRH	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CAAGAGGCGCTCACCTCC GGG (SEQ ID NO: 1686)	TATGAACCACTAAGGCGTcaacttgtgCAAGAGGCGCTCAC (SEQ ID NO: 3259)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCCCGAG GTGAGGCGCGTCTGCAAGAAGGTGGAACGCCTTAGTGGTTCATA (SEQ ID NO: 4332)	SHRT6_-_4162479 23-P1P2	SHRT6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGAATGTCCGTTGAATTA GCG (SEQ ID NO: 1687)	TATGAACCACTAAGGCGTcaacttgtgGAAGAATTGTCGGTTG (SEQ ID NO: 3260)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCGTAATTC ACCGACATCCTCCAAGAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4333)	SHRT6_+_41825177 23-P1P2	SHRT6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCAAGTTGCGCCTCC CGG (SEQ ID NO: 1688)	TATGAACCACTAAGGCGTcaacttgtgGGGCAAGTTGCGCC (SEQ ID NO: 3261)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCGGGAG GCCGCACTTGCCCACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4334)	SHRT6_+_41824692 23-P1P2	SHRT6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGACAGAGCGACTTGG TGA (SEQ ID NO: 1689)	TATGAACCACTAAGGCGTcaacttgtgGAGACAGAGCGAC (SEQ ID NO: 3262)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCACCAAG TCGCTCTGTCTCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4335)	SKIV2L_-_31926997 2-P1P2	SKIV2L	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCCACGTGCCCGAAGG GGT (SEQ ID NO: 1690)	TATGAACCACTAAGGCGTcaacttgtgGCTCCACGTGCCC (SEQ ID NO: 3263)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACACCCTTC GGGCCGTCGAAGCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4336)	SKIV2L_+_31927098 23-P1P2	SKIV2L	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGATGGAGACAGAGCG GCT (SEQ ID NO: 1691)	TATGAACCACTAAGGCGTcaacttgtgGTGATGGAGACAG AGGCGACTgcttaagagctaagctgtGGTAAAGTCTGCTGGTGA (SEQ ID NO: 3264)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACAGTCGCT CTTGTCTCCAATCCACCAAGAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4337)	SKIV2L_-_31926981 2-P1P2	SKIV2L	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACAGGGAGAGGGGCAG TAC (SEQ ID NO: 1692)	TATGAACCACTAAGGCGTcaacttgtgGACAGGAGAGAGGGG (SEQ ID NO: 3265)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGGTAGGTG CCCCTCTCCTGTCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4338)	SLX1A_-_30205265 2-3-ENST0000002	SLX1A	ENST0000 0025133 6	SLX1B	SLX1B_-_29465914 2-3-ENST0000	ENST0000 0033018 5	BOLA2	ENST0000 0033097 3	BOLA2B	BOLA2B_-_30205265 2-3-ENST0000	BOLA2B	ENST0000 0030527 4
GGACTCTCTGGGGGCACT CGG (SEQ ID NO: 1693)	TATGAAGCACTAAGGCGTcaacttgtgGGCTCTCCGGGCAC (SEQ ID NO: 3266)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCGGATA TCCCCGGAGAAGCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4339)	SLX1A_+_30205279 23-ENST0000002 5133.6	SLX1A	ENST0000 0025133 5	SLX1B	SLX1B_+_29466937 23-ENST0000003 0018.5	ENST0000 0033018 5	BOLA2_+_29466937 23-ENST0000003 0018.3	ENST0000 0033097 8	BOLA2B	BOLA2B_+_30205279 23-ENST0000003 0527.4	BOLA2B	ENST0000 0030527 4
GCAAAGCTCTCGGGGGACT ATC (SEQ ID NO: 1694)	TATGAACCACTAAGGCGTcaacttgtgGCAAAGCTCTCGGGG (SEQ ID NO: 3267)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAAGGATAGTCC GACTATCGttaagagctaagctgtCGTAAAGTCTGCTGGTGCA (SEQ ID NO: 4340)	SLX1A_+_30205282 23-ENST0000002 5133.6	SLX1A	ENST0000 0025133 6	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGTGTAGAAGTCCGA CCG (SEQ ID NO: 1695)	TATGAACCACTAAGGCGTcaacttgtgGGGGTGTAGAAGT CCGACCGttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3268)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCGGGTCGG ACTTCTACAGCCCCAAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4341)	SLX1B_+_29465983 23-ENST0000003 0018.5	SLX1B	ENST0000 0033018 5	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTGCCCCTCTCCTGTCT GG (SEQ ID NO: 1696)	TATGAAGCACTAAGGCGTcaacttgtgGGTGCCCCTCTCCT (SEQ ID NO: 3269)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAAGCCCAGACA GGAAGAGGGGCACCCCAAACAAGGTTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4342)	SLX1B_+_29465866 23-ENST0000003 0018.5	SLX1B	ENST0000 0033018 5	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCGGCGGCGCCGCGG AGCGT (SEQ ID NO: 1697)	TATGAACCACTAAGGCGTcaacttgtgGCCGGCGGCGCCGCGG (SEQ ID NO: 3270)	TGACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACGACCTCCG CGCGCCGCGGCGCCGCGGCAACAAGGTTGGACGCCTTAGTGGTTCA A (SEQ ID NO: 4343)	SLX4_+_3661556 23-P1P2	SLX4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGTTAAGAAGGTGCCG CCG (SEQ ID NO: 1698)	TATGAAGCACTAAGGCGTcaacttgtgGCCGTTAAGAAGT (SEQ ID NO: 3271)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCGCGCC ACCTTCTTAAGCGCCAACAAGGTGCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4344)	SLX4_+_3661589 23-P1P2	SLX4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGGGACTCCGCGACCC GGG (SEQ ID NO: 1699)	TATGAAGCACTAAGGCGTcaacttgtgGCTGGGACTCCGG (SEQ ID NO: 3272)	TCAGCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACCCCGGGGT GCCGACTCCCCACAAGCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4345)	SLX4_+_3661515 23-P1P2	SLX4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCAGAGTACAGAGAAC AGA (SEQ ID NO: 1700)	TATGAACCACTAAGGCGTcaacttgtgCCCAGAGTACAG AAACACAGcttaagagctaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3273)	TCACCAGCAGAGCTTTACGCAGCTTAGCTCTTAAACTCTGTTTC TGTATCTCTGGGGCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4346)	SMARCA2_+_20165334 23-P1P2	SMARCA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTACTCTGGGTCAGTCA GAG (SEQ ID NO: 1701)	TATGAAGCACTAAG3GCTcaacctgtgG7ACTCTGGGTGA cACCCAGAGTtaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3274)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAAGCCTCTGAGT A (SEQ ID NO: 4947)	SMARCA2_20153833:P1P2	SMARCA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GTTGCTCGCGAGGAGTG TGG (SEQ ID NO: 1702)	TATGAACCACTAAGGCTcaacctgtgG7TGCTCGCGAGG AGTCTGTGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3275)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAAGCACTC CTCGCGAGGAACCAAGAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4948)	SMARCA2_+20153812:P1P2	SMARCA2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
CGCCGCCGCCGCCAAAAC GAG (SEQ ID NO: 1703)	TATGAACCACTAAGCGCTcaacctgtgG3CCGCCGCCGCGCC fAAACGAGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3276)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACCTCGTTTG GCGCGCGGCGCGCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4949)	SMARCA4_+11077242:3-P1P2	SMARCA4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTTCGCCTCGGTTTGGC CGG (SEQ ID NO: 1704)	TATGAAGCACTAAGGCTcaacctgtgCGCTTCCCCTCGTT TGGCGGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3277)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAAGCCGCCAA ACGAGGCGGAAGCCCAAGAAGGTGGAGCGCCTTAGTGGTTCAT A (SEQ ID NO: 4950)	SMARCA4_-11077412:3-P1P2	SMARCA4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGAGAGGCGGGGAAGT CGA (SEQ ID NO: 1705)	TATGAACCACTAAGGCTcaacctgtgGggagagcgcggAAG TCGAgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3278)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACTCGACTTC CCGCGCGCCTCCCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4951)	SMARCA4_-11071834:2-P1P2	SMARCA4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGTTGAGGTAAGCGC CGG (SEQ ID NO: 1706)	TATGAACCACTAAGGCTcaacctgtgGGCCGTTGAGGTAA GCGCCGGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3279)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACCCGGCGC TTACCTCACGGGCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4952)	SMARCA5_-144346513:23-P1P2	SMARCA5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTGTGAACTATGAGCC CCG (SEQ ID NO: 1707)	TATGAACCACTAAGGCTcaacctgtgCGGGTGTGAACCTAG AGCCCCGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3280)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACCCGGCGCT CTATGGTTCCACCCCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4953)	SMARCA5_-144346778:23-P1P2	SMARCA5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGGAAGGCGCGCTCTC GGG (SEQ ID NO: 1708)	TATGAACCACTAAGGCTcaacctgtgGTGGAAGGCGCGCTC TCTCGGGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3281)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACCCCGAGA GCGGCGCGTTCCACCAACAAGGTGGAGCGCCTTAGTGGTTCATA (SEQ ID NO: 4954)	SMARCA5_+144435110:23-P1P2	SMARCA5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGAGGTTGCAGTGCATG CTA (SEQ ID NO: 1709)	TATGAAGCACTAAGGCGTcaacctgtgGAGAGGTTGCAGTG CATGCTAgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3282)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACTAGCATGC ACTCCACCCTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4955)	SMARCAD1_+_96128831:23-P1P2	SMARCAD	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GATCCGCCGCCGAAGTTC ACG (SEQ ID NO: 1710)	TATGAAGCACTAAGGCTcaacctgtgGATCCGCCGCCGAAA GTTGACGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3283)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACCCGTCAACT TCCGGGCGGAATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4956)	SMARCAD1_+_96128011:23-P1P2	SMARCAD	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACGCTGAGAGGAAGCT GAT (SEQ ID NO: 1711)	TATGAACCACTAAGGCTcaacctgtgGCAGCTGAGAGGAG AAGCTGATgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3284)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACATCAGCGT CCTCTCAGCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4957)	SMARCAD1_+_96128683:23-P1P2	SMARCAD	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGCGTGGCCGCTACA ATA (SEQ ID NO: 1712)	TATGAAGCACTAAGGCGTcaacctgtgGAGGCGTGGCCGC TACAATAgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3285)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACTATTGTAG CGGCGCAGGCGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4958)	SMARCAL1_-21727235:23-P1P2	SMARCAL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GATCCAAGTGGAGATCG TAG (SEQ ID NO: 1713)	TATGAAGCACTAAGGCGTcaacctgtgGATCGAAGTGCGAG ATCGTAGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3286)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAAGCTACGATC TCCACTTGGATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4959)	SMARCAL1_-21727482:23-P1P2	SMARCAL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGTAGCGCCGCAGGCCT GAA (SEQ ID NO: 1714)	TATGAAGCACTAAGGCGTcaacctgtgGTGTAGCGCCGCAG GCCTGAAgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3287)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAAGTTCAGGGC CTGCGCCGCTACACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4960)	SMARCAL1_+_21727720:23-P1P2	SMARCAL1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
CGCCGCCGCGAGACGACG ACC (SEQ ID NO: 1715)	TATGAAGCACTAAGGCGTcaacctgtgAGACGCGGCGACGAC GAGACCgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3288)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACGGTCGTC GTCTCGCGGCGCGCCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4961)	SMARCB1_+_24129177:23-P1	SMARCB1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
GTACTTGGTCGTCGTCTCG CCG (SEQ ID NO: 1716)	TATGAAGCACTAAGTACGTcaacctgtgCGCCTGGTCGTCG TCTGCGGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3289)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACCCGCAGCA CGAGCGACCGACGCCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4962)	SMARCB1_-24129197:23-P1	SMARCB1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na
CGCCGTCGTCGTCGTCGT CTG (SEQ ID NO: 1717)	TATGAACCACTAAGGCGTcaacctgtgCGTCGTCGTCGTC GTCTGTGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3290)	TCACCAGCAGAGTTTACGCAGCTTAGCTCTTAAACCACGACGA CGAGCCAGACGCCGCCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4963)	SMARCB1_-24129194:23-P1	SMARCB1	P1	na	na	na	na	na	na	na	na	na	na	na	na	na

GAGAAGCTTCGGAACATAC GG (SEQ ID NO: 1718)	TATGAAGCACTAAGGCGTcaacctgtgGAGAACGTCGGAA	CATACGGgttaagagtcaagtgGTAAAGTCTGCTGGTGA	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCCGTATGT (SEQ ID NO: 4964)	TCCGAGGTTCTCCAACAAGGTCGACGCCTTAGTGGTTCATA	SMARCB1_+ 24133853 2 3-P2	SMARCB1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTAGTCGCCTCCAGAG TGA (SEQ ID NO: 1719)	TATGAACCACTAAGGCGTcaacctgtgGCTAGTCGCCTCCA	GAGTGAgttaagagtcaagtgGTAAAGTCTGCTGGTGA	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTCACTCTG (SEQ ID NO: 4965)	GAGTCGACTTAGTGGTTCATA	SMARCB1_+ 24133860 2 3-P2	SMARCB1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAAGATACCCCTCACTCT GG (SEQ ID NO: 1720)	TATGAACCACTAAGGCGTcaacctgtgGAAGATACCCCTCA	CTCTTGAgttaagagtcaagtgGTAAAGTCTGCTGGTGA	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCCACGAGT (SEQ ID NO: 4966)	GAGGGGTTATCTCCCAACAAGGTCGACGCCTTAGTGGTTCATA	SMARCB1_- 24134004 2 3-P2	SMARCB1	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGCCACGAACGGACAA CAA (SEQ ID NO: 1721)	TATGAACCACTAAGGCGTcaacctgtgCACGCCACGAGGGT	ACCACCAAgttaagagtcaagtgGTAAAGTCTGCTGGTGA	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTTTGTTCTC (SEQ ID NO: 4967)	CACTCGTTGGCTCCAACAAGGTCGACGCCTTAGTCGTTCATA	SMARCC1_+ 47822975 2 3-P1P2	SMARCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCAGGGTTGTGGGGAC GAC (SEQ ID NO: 1722)	TATGAACCACTAAGGCGTcaacctgtgGGGCAGGGTTGTGGGGA	aagttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3295)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGCTCTGCTC (SEQ ID NO: 4968)	CCACAAGCCTTGTGCCCAACAAGGTCGACGCCTTAGTGGTTCATA	SMARCC1_+ 47823295 2 3-P1P2	SMARCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTCGCGCGCGGGAAC GACG (SEQ ID NO: 1723)	TATGAACCACTAAGGCGTcaacctgtgGggtgcgcgcgggaacga	ccgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3296)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGGTCGTT (SEQ ID NO: 4969)	CCCGCGGCGACCCCAACAAGGTCGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4969)	SMARCC1_+ 47823940 2 3-P1P2	SMARCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGTTGTCGAACTCGGG TCA (SEQ ID NO: 1724)	TATGAACCACTAAGGCGTcaacctgtgGACGTTGTCGAACT	GGGTCAgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3297)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACTGACCCA (SEQ ID NO: 4970)	GTTCGACAAACCGTCCAACAAGGTTGGACGCCTTAGTGGTTCATA	SMARCC2_- 58583185 2 3-P1P2	SMARCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGAAGCCGGAAGAAGA TGG (SEQ ID NO: 1725)	TATGAACCACTAAGGCGTcaacctgtgGaccccagccgggaagaat	gaggttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3298)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCCATCTTC (SEQ ID NO: 4971)	TCCCGGCTCGGGCCAACAAGGTTGGACGCCTTAGTGGTTCATA	SMARCC2_+ 56583238 2 3-P1P2	SMARCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGCGGGAGACAAAGGGA TCG (SEQ ID NO: 1726)	TATGAACCACTAAGGCGTcaacctgtgGgggcgggAACAAAGG	GAGCGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3299)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGGTCGCC (SEQ ID NO: 4972)	TTTGTCCGCGCCCCAACAAGGTCGACGCCTTAGTGGTTCATA	SMARCC2_+ 56583078 2 3-P1P2	SMARCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCAGGGGCACGGAGAC TAT (SEQ ID NO: 1727)	TATGAAGCACTAAGGCGTcaacctgtgGCTCAGGGGCAGGG	AGAGTAAgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3300)	TCAGCACGCAGACTTTACGCAGCTTAGCTCTTAAACATACTCTC (SEQ ID NO: 4973)	GACTGCCTGAGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4973)	SMARCD1_+ 50479015 2 3-P1P2	SMARCD1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGATAGAGCGGCGGCGAC TAG (SEQ ID NO: 1728)	TATGAACCACTAAGGCGTcaacctgtgGGATAGAGCGCGGGG	CGACTAGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3301)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGCTAGTGG (SEQ ID NO: 4974)	CCCCCGCTCTATCCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4974)	SMARCD1_+ 50479394 2 3-P1P2	SMARCD1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGCGGCCGGCCGGCGA GGGG (SEQ ID NO: 1729)	TATGAACCACTAAGGCGTcaacctgtgGgcgggcggccggcgac	ggagttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3302)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGCCCGTGC (SEQ ID NO: 4975)	CGCGCGTGCCGGGCCCAACAAGGTCGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4975)	SMARCD1_+ 50479937 2 3-P1P2	SMARCD1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGCGGGGAGTCCGGGA GTC (SEQ ID NO: 1730)	TATGAAGCACTAAGGCGTcaacctgtgCGAGCGGCGAGCGCGAG	TCTCGGTCCCGTTCCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4976)	TCAGCACGCAGACTTTACGCAGCTTAGCTCTTAAACGCACATCG (SEQ ID NO: 4976)	CTCGGTCCCCGTTCCCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4976)	SMARCD2_+ 51920028 2 3-P1P2	SMARCD2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACGCGGGGAGTCCGGGA TCG (SEQ ID NO: 1731)	TATGAACCACTAAGGCGTcaacctgtgGGACGCCCGGAGTC	CCGGATCTgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3304)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGCGATCCG (SEQ ID NO: 4977)	GACTCGCCGGCTCCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4977)	SMARCD2_+ 51920355 2 3-P1P2	SMARCD2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCACCGGAGTGTGGACTCG ACG (SEQ ID NO: 1732)	TATGAACCACTAAGGCGTcaacctgtgAGCACCGGAGTGTGGA	CTCGACGgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3305)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACCGTCGAG (SEQ ID NO: 4978)	TCCACTCCCGTTGCCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 4978)	SMARCD2_+ 51922077 2 3-P1P2	SMARCD2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAAGCTCCGCGACGCGAG GGG (SEQ ID NO: 1733)	TATGAAGCACTAAGGCGTcaacctgtgGAAGCTCCGCGACG	CGTCCGGAgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3306)	TCAGCACGCAGACTTTACGCAGCTTAGCTCTTAAACCCCCTCG (SEQ ID NO: 4979)	CGTCCGGAGCGTTCCAACAAGGTTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4979)	SMARCE1_+ 39804093 2 3-P1P2	SMARCE1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTTGGAAACACTCACCC GC (SEQ ID NO: 1734)	TATGAACCACTAAGGCGTcaacctgtgGCTTGGAAACACTC	ACCCGCgttaagagtcaagtgGTAAAGTCTGCTGGTGA (SEQ ID NO: 3307)	TCACGACGCAGACTTTACGCAGCTTAGCTCTTAAACGGGTGTG (SEQ ID NO: 4980)	AGTGTTTCCAAGCCCAACAAGGTCGACGCCTTAGTGGTTCATA	SMARCE1_- 39804033 2 3-P1P2	SMARCE1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGGGCGGAGCGGGGGA CAAA (SEQ ID NO: 1735)	TATGAACCACTAAGGCGTcaacctgtgGGGGGGTGAGCGC GGGACAAAgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3308)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACTTTGTCCC 1 GCCGTCGCCGCCCAAOAAGGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 4981)	SMARCE1_+ 38804053 2 3-P1P2	SMARCE1	P1P2	na	na	na	na	na	na	na	na	na	na	na
GGCCGTACGCCCGAGAA CTG (SEQ ID NO: 1736)	TATGAACCACTAAGGCGTcaacctgtgGGCCGTTAGCTCTTA AGGGCGTACGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4982)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGACTTCTC 2 GGGCGTACGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4982)	SMC1A_- 53449605 2 3-P1P2	SMC1A	P1P2	RBC1_- 53449606 2 3-P1P2	RBC1	P1P2	na	na	na	na	na	na	na	na
GGAACTGAGGTAGCCCG CGC (SEQ ID NO: 1737)	TATGAACCACTAAGGCGTcaacctgtgGAAACTGAGGTAG CTACCTCAGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4983)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGACGTCGCG 2 CTACCTCAGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4983)	SMC1A_- 53449619 2 3-P1P2	SMC1A	P1P2	na	na	na	na	na	na	na	na	na	na	na
GGCCGCGGTGCGGCAGTAG GAC (SEQ ID NO: 1738)	TATGAACCACTAAGGCGTcaacctgtgCGCGCCCGGCGGCA GTACGGACGTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4984)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGTCGTACT 2 GCCGCCGGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4984)	SMC1A_- 53449660 2 3-P1P2	SMC1A	P1P2	na	na	na	na	na	na	na	na	na	na	na
GCGCGGGGCGCTTGATAA CGC (SEQ ID NO: 1739)	TATGAACCACTAAGGCGTcaacctgtgGCGCGGGGCTTGATC ATAACCGcgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3312)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGCGCTTATC 1 AACCGCCCCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4985)	SMC1B_+_4 8809459 23-P1P2	SMC1B	P1P2	na	na	na	na	na	na	na	na	na	na	na
GTAACGCGGGTGAGGCG TGG (SEQ ID NO: 1740)	TATGAACCACTAAGGCGTcaacctgtgGTAACGCGGGTGA GCGCGTGGgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3313)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCCACGCC 1 TCACCGCGCGTTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4986)	SMC1B_+_4 9809456 23-P1P2	SMC1B	P1P2	na	na	na	na	na	na	na	na	na	na	na
GTGATAACGCGGTGTGAG GCG (SEQ ID NO: 1741)	TATGAACCACTAAGGCGTcaacctgtgGTGATAACGCGGG TGAGCGCGgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3314)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGCGCTCA 1 CCCGCGCTTATCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4987)	SMC1B_+_4 5809459 23-P1P2	SMC1B	P1P2	na	na	na	na	na	na	na	na	na	na	na
GAAATCTCTGGCCCGGCCC AAC (SEQ ID NO: 1742)	TATGAACCACTAAGGCGTcaacctgtgGAAATCTCTGGCCCG GCCCCAACgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3315)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGTTGGGCC 2 CGGGGCCAACTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4988)	SMC2_+_10 8856781 23-P1P2	SMC2	P1P2	na	na	na	na	na	na	na	na	na	na	na
GGCGGGGTGTGTTGAGAGCG GTG (SEQ ID NO: 1743)	TATGAACCACTAAGGCGTcaacctgtgGGCGGGGTGTTGAG AGCGGGTGgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3316)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACCACCGCT 2 CTCAACACCCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4989)	SMC2_- 10685683 23-P1P2	SMC2	P1P2	na	na	na	na	na	na	na	na	na	na	na
GGTGCGTGAGCACGCGCC GGT (SEQ ID NO: 1744)	TATGAAGCACTAAGGCGTcaacctgtgCGTGAGCACGCGC GGTGCGTCgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3317)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACACCGTGCGC 2 GTGCTCACGCACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4990)	SMC2_- 10695788 23-P1P2	SMC2	P1P2	na	na	na	na	na	na	na	na	na	na	na
GCAGCAACCCGCACGCGCC GCA (SEQ ID NO: 1745)	TATGAACCACTAAGGCGTcaacctgtgCAGCAACCCGCAC GCGCCGCAgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3318)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAAGTGCGGGCC 1 GTGCGGTTGGTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4991)	SMC3_+_11 2327512 23-P1P2	SMC3	P1P2	na	na	na	na	na	na	na	na	na	na	na
GCCCCGGGAGCAGCAACC AACA (SEQ ID NO: 1746)	TATGAACCACTAAGGCGTcaacctgtgGCCCCGGGAGCAGC AACAGGCAgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3319)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACTGCGGTT 1 GCTGCTCCGGGGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 4992)	SMC3_+_11 2327520 23-P1P2	SMC3	P1P2	na	na	na	na	na	na	na	na	na	na	na
GGTCCAGCCTGTCTCGAGA CAG (SEQ ID NO: 1747)	TATGAAGCACTAAGGCGTcaacctgtgGGTCCAGCCTGTGT GAGACACAGgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3320)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAAGCTGTCTCC 1 ACAAGGCTGTGAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4993)	SMC3_+_11 2127677 23-P1P2	SMC3	P1P2	na	na	na	na	na	na	na	na	na	na	na
GGCGGAGCCGAAAGACCC GGT (SEQ ID NO: 1748)	TATGAACCACTAAGGCGTcaacctgtgGCCGGAAAGACCC GGTCGGTCgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3321)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGACCGGTG 2 TTTCGGCTCCGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4994)	SMC4_- 1601174752 3-P1	SMC4	P1	na	na	na	na	na	na	na	na	na	na	na
GCCGGTAGGAGCGGCGGA GGT (SEQ ID NO: 1749)	TATGAACCACTAAGGCGTcaacctgtgGCCGGTAGGAGCGG TGGAAGGTgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3322)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACACCTCCCG 2 GCTCCTACCGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4995)	SMC4_- 1601174892 3-P1	SMC4	P1	na	na	na	na	na	na	na	na	na	na	na
GGGGACCTCGAGCTCGAAA CAC (SEQ ID NO: 1750)	TATGAAGCACTAAGGCGTcaacctgtgGGGACCTCGAGCTC GAAACACAgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3323)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACGTTCTCG 2 GCTCCGGGGCTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4996)	SMC4_- 1601174712 3-P1	SMC4	P1	na	na	na	na	na	na	na	na	na	na	na
GAATTTCTATGTATCTCAG AA (SEQ ID NO: 1751)	TATGAACCACTAAGGCGTcaacctgtgGAATTTCTATGTAT CCACGAAgttaagagtaagctgGTAAAGTCTGTGGTGA (SEQ ID NO: 3324)	TCACCAGCAGAGACTTTAGCGAGGTTAGCTCTTAAACTTCTTCGAT 2 ACATACGAAAATTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4997)	SMC4_- 1601221693 3-P2	SMC4	P2	na	na	na	na	na	na	na	na	na	na	na

GATGGTCCAAGTCAATTCCA (SEQ ID NO: 1752)	TATGAACCACTAAGGCGTccaccttgttgGATGGTCCAAGTCAATTCCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3325)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGAATTGACTTGGACCATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4898)	2	SMC4_+_160122249.23-P2	SMC4	P2	na	na	na	na	na	na	na	na	na
GCTTCTGTCTATCACATAAG (SEQ ID NO: 1753)	TATGAACCACTAAGGCGTccaccttgttgGCTTCTGTCTATCACATAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3326)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTATGTGATAGACAGAAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4899)	2	SMC4_-_160122206.23-P2	SMC4	P2	na	na	na	na	na	na	na	na	na
GCGGCGGCACCAAGTATCCG (SEQ ID NO: 1754)	TATGAACCACTAAGGCGTccaccttgttgGcggcggcACCAAGTATCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3327)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGATACTTGGTGCCGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4900)	1	SMG1_+_18937294.23-P1P2	SMG1	P1P2	na	na	na	na	na	na	na	na	na
GGCCAGTCATTCCAGCTCCG (SEQ ID NO: 1755)	TATGAACCACTAAGGCGTccaccttgttgGGCCAGTCATTCCAGCTCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3328)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGAGCTGGAATGACTGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4901)	1	SMG1_-_18937299.23-P1P2	SMG1	P1P2	na	na	na	na	na	na	na	na	na
GTTCCAGCTCCGCGGATACT (SEQ ID NO: 1756)	TATGAACCACTAAGGCGTccaccttgttgGTTCCAGCTCCGCGGATACTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3329)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGTATCCGCGGAGCTGGAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4902)	1	SMG1_-_18937307.23-P1P2	SMG1	P1P2	na	na	na	na	na	na	na	na	na
GAAGGAACCGGAAACGGGAT (SEQ ID NO: 1757)	TATGAACCACTAAGGCGTccaccttgttgGAAGGAACCGGAAACGGGATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3330)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATCCCGTTTCCGGTTCCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4903)	1	SMUG1_+_54582746.23-P1P2	SMUG1	P1P2	na	na	na	na	na	na	na	na	na
GGGACCAGGTGAGGAAGTAG (SEQ ID NO: 1758)	TATGAACCACTAAGGCGTccaccttgttgGGGACCAGGTGAGGAAGTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3331)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTACTTCCTCACCTGGTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4904)	1	SMUG1_+_54582719.23-P1P2	SMUG1	P1P2	na	na	na	na	na	na	na	na	na
GTGGGAATCCGGGACTGTGG (SEQ ID NO: 1759)	TATGAACCACTAAGGCGTccaccttgttgGTGGGAATCCGGGACTGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3332)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACAGTCCCGGATTCCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4905)	1	SMUG1_+_54582695.23-P1P2	SMUG1	P1P2	na	na	na	na	na	na	na	na	na
GAAGATGGCGCTCACCAGGT (SEQ ID NO: 1760)	TATGAACCACTAAGGCGTccaccttgttgGAAGATGGCGCTCACCAGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3333)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCTGGTGAGCGCCATCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4906)	2	SNW1_+_78227451.23-P1	SNW1	P1	C14orf178_+_78227451.23-P1P2	C14orf178	P1P2	na	na	na	na	na	na
GTCGCTCGCGCTGGAAGAAG (SEQ ID NO: 1761)	TATGAACCACTAAGGCGTccaccttgttgGTCGCTCGCGCTGGAAGAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3334)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTCTTCCAGCGCGAGCGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4907)	2	SNW1_+_78227476.23-P1	SNW1	P1	C14orf178_+_78227476.23-P1P2	C14orf178	P1P2	na	na	na	na	na	na
GGAGCGGGTGAACTAAACCC (SEQ ID NO: 1762)	TATGAACCACTAAGGCGTccaccttgttgGGAGCGGGTGAACTAAACCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3335)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGTTTAGTTCACCCGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4908)	2	SNW1_-_78227263.23-P1	SNW1	P1	na	na	na	na	na	na	na	na	na
GCCCCGTACGGATACCGGAA (SEQ ID NO: 1763)	TATGAACCACTAAGGCGTccaccttgttgGCCCCGTACGGATACCGGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3336)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCCGGTATCCGTACGGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4909)	2	SNW1_+_78221331.23-P2	SNW1	P2	na	na	na	na	na	na	na	na	na
GCCGGAAAGGCTGGATACCT (SEQ ID NO: 1764)	TATGAACCACTAAGGCGTccaccttgttgGCCGGAAAGGCTGGATACCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3337)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGGTATCCAGCCTTTCCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4910)	2	SNW1_+_78221318.23-P2	SNW1	P2	na	na	na	na	na	na	na	na	na
GGAAGAGAACCTCCCCGTA (SEQ ID NO: 1765)	TATGAACCACTAAGGCGTccaccttgttgGGAAGAGAACCTCCCCCGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3338)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACGGGGGAGGTTCTCTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4911)	2	SNW1_+_78221343.23-P2	SNW1	P2	na	na	na	na	na	na	na	na	na
GAAGGGCGCAGCCTAGGACA (SEQ ID NO: 1766)	TATGAACCACTAAGGCGTccaccttgttgGAAGGGCGCAGCCTAGGACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3339)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTCCTAGGCTGCGCCCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4912)	1	SPO11_-_55904890.23-P1P2	SPO11	P1P2	na	na	na	na	na	na	na	na	na
GACGAGGCCTCGGGCCCCAT (SEQ ID NO: 1767)	TATGAACCACTAAGGCGTccaccttgttgGACGAGGCCTCGGGCCCCATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3340)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATGGGGCCCGAGGCCTCGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4913)	1	SPO11_+_55904935.23-P1P2	SPO11	P1P2	na	na	na	na	na	na	na	na	na
GGCACGCAGCCACGCCCCAA (SEQ ID NO: 1768)	TATGAACCACTAAGGCGTccaccttgttgGGCACGCAGCCACGCCCCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3341)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGGGGCGTGGCTGCGTGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4914)	1	SPO11_-_55904873.23-P1P2	SPO11	P1P2	na	na	na	na	na	na	na	na	na

GAGCCCAAGCCGGCTAG CCG (SEQ ID NO: 1769)	TATGAACCACTAAG3GCTcaacctgtgGACCCCAAGCCGC GCCGGTTGG3G3TCCAACAAG3TGGACGGCTTAGTGTGGTGA (SEQ ID NO: 3342)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCCGGCTAG CGTCGTAGGAGATCCGAACAAG3TGGACGGCTTAGTGTGGTGA (SEQ ID NO: 4915)	1	SPRTN_+_2 31474047 23-P1P2	SPRTN	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATCTCGTAGGAGCTAGT CC (SEQ ID NO: 1770)	TATGAACCACTAAG3GCTcaacctgtgGATCTCCTAGGAG CTAGTCCgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3343)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCCGGCTAG CTCGTAGGAGATCCGAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4916)	1	SPRTN_-_ 23147464 5 23-P1P2	SPRTN	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGATGCCGGCTCTCTCG AGA (SEQ ID NO: 1771)	TATGAACCACTAAG3GCTcaacctgtgGAGATCCCGGCT CTCGAG3gttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3344)	TCAACCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCTCTCGAGAT GCCGGCATCTCCCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4917)	1	SPRTN_+_2 31474057 23-P1P2	SPRTN	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGCATTTCCGCGCCTC CTG (SEQ ID NO: 1772)	TATGAACCACTAAG3GCTcaacctgtgCATCCATTTCCGCGG CCTCCTCTGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3345)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCAGGAGCC CCCGGGAAATGCTCCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4918)	1	SRCAP_-_ 30710677 2 3-P1P2	SRCAP	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGATTCCGAAGCGAGTGG TGG (SEQ ID NO: 1773)	TATGAACCACTAAG3GCTcaacctgtgGGATTCCGAAGCGA gtgtgggttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3346)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCCACCACT CGGTCGG3AATCCCAACAAG3TGGACGCCGTTAGTGGTTCATA (SEQ ID NO: 4919)	1	SRCAP_-_ 30709514 2 3-P1P2	SRCAP	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCAGTCCGGTCGGGAGG 3GT (SEQ ID NO: 1774)	TATGAACCACTAAG3GCTcaacctgtgGTCAGTCCGTCG3GA GAG3GGTgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3347)	TCAACCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCAGCCCTC CCGACGGACTGACCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4920)	1	SRCAP_-_ 30710513 2 3-P1P2	SRCAP	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
3CGGCAGAGCGACGCCG CAG (SEQ ID NO: 1775)	TATGAACCACTAAG3GCTcaacctgtgCGGCGGCAGAG3GAG CCCGGCAGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3348)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCTCGCGGG CTCGGCTCTGGCGCCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4921)	2	SRSF1_-_ 58084574 2 3-P1P2	SRSF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCTCGGCGCACGGATA TGG (SEQ ID NO: 1776)	TATGAACCACTAAG3GCTcaacctgtgGGCTCGGCGCACG GATATCGGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3349)	TCACCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCCATATCC GATATCGGCGACGCCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4922)	2	SRSF1_-_ 56084638 2 3-P1P2	SRSF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGACGTGGTGCCGC TGC (SEQ ID NO: 1777)	TATGAAGCACTAAG3GCTcaacctgtgGGGAAGACGTGGTG CCCGCTGCgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3350)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACGCAGCGCG CACGACGTCTCCCCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4923)	2	SRSF1_+_56 084585 23-P1P2	SRSF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCGTGAACGCGCAAAAC TGG (SEQ ID NO: 1778)	TATGAAGCACTAAG3GCTcaacctgtgGCGCGTGAACGCGC AAAACTGGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3351)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCCAG3TTTG CCGTTCACCCG3CCAACAAGGTG3ACGGCTTAGTGGTTCATA (SEQ ID NO: 4924)	1	SS18_-_ 23670693 2 3-P1P2	SS18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTTGCCGATAGTGTGG TGA (SEQ ID NO: 1779)	TATGAACCACTAAG3GCTcaacctgtgGGACTTGCCGATA GTG3GTGAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3352)	TCACCAGCAGACTTTACGCAGCTTAAGCTCGTTAAAGTCACCACT ATCGGCGAAGTCCCAACAAG3TGGACGCCGTTAGTGGTTCATA (SEQ ID NO: 4925)	1	SS18_+_236 70540 23-P1P2	SS18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTCGGGACTTGCCGA TAG (SEQ ID NO: 1780)	TATGAACCACTAAG3GCTcaacctgtgGGCTCGGGACTTG CCGATAGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3353)	TCACCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCTATCGG CAAAGTCCCGAGCCCACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4926)	1	SS18_+_238 70548 23-P1P2	SS18	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGCTCG3G3TTAGATC GTC (SEQ ID NO: 1781)	TATGAAGCACTAAG3GCTcaacctgtgCGCTGCTCCGCGTTA GATCG3TCgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3354)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACG3AGGATCT AAGCCCGAAG3CAGCCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4927)	2	SSBP1_-_ 14143382 10 23-P1P2	SSBP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGACACCCTTTACCG3AA TA (SEQ ID NO: 1782)	TATGAACCACTAAG3GCTcaacctgtgG3G3ACACCCTTTAC CGAAGAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3355)	TCAACCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACCTTAC AAAAGG3TGTCCCCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4928)	2	SSBP1_+_14 14383842 23-P1P2	SSBP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGTTAGATCGTCAGG TGA (SEQ ID NO: 1783)	TATGAACCACTAAG3GCTcaacctgtgG3G3GTTAGATCGT CAGGGTGAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3356)	TCAACCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACTCACGTGA CGATCTAACCCCCAACAAG3TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 4929)	2	SSBP1_-_ 14143382 16 23-P1P2	SSBP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
G3CAG3CGCG3AAAAG3CTT CCC (SEQ ID NO: 1784)	TATGAAGCACTAAG3GCTcaacctgtgCGAC3CGG3G3AAAA TTTTCCCGCGgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3357)	TCAGCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACG3AGCAAG3C TTTTCCCGCGGTGCCAACAAG3TGGACGCCGTTAGTGGTTCATA (SEQ ID NO: 4930)	2	SSRP1_+_57 03286 23-P1P2	SSRP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCTCTGTGGCCTGT GTA (SEQ ID NO: 1785)	TATGAACCACTAAG3GCTcaacctgtgGCCTCTCTGTGGC CTGTGTAgttaagagctaagtgtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3358)	TCACCAGCAGACTTTACGCAGCTTAAGCTCGTTAAACTACACAG GCCACAG3AGCAG3GCCAACAAG3TGGACGGCTTAGTGGTTCAT A (SEQ ID NO: 4931)	2	SSRP1_+_57 03345 23-P1P2	SSRP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGCCTGTGTACGGCTTCCGG (SEQ ID NO: 1786)	TATGAACCACTAAGGCGTccaccttgttgGGCCTGTGTACGGCTTCCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3359)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGAAGCCGTACACAGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4932)	2	SSRP1_+_57103335.23-P1P2	SSRP1	P1P2	na	na	na	na	na	na	na	na	na
GATCCTAACCCGCCATCTGG (SEQ ID NO: 1787)	TATGAACCACTAAGGCGTccaccttgttgGATCCTAACCCGCCATCTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3360)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCAGATGGCGGGTTAGGATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4933)	1	STAG1_-_136471201.23-P1P2	STAG1	P1P2	na	na	na	na	na	na	na	na	na
GCACCGCGGCCAGCAGAGGG (SEQ ID NO: 1788)	TATGAACCACTAAGGCGTccaccttgttgGCACCGCGGCCAGCAGAGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3361)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCTCTGCTGGCCGCGGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4934)	1	STAG1_+_136471246.23-P1P2	STAG1	P1P2	na	na	na	na	na	na	na	na	na
GTGCCACCAGATGGCGGGTT (SEQ ID NO: 1789)	TATGAACCACTAAGGCGTccaccttgttgGTGCCACCAGATGGCGGGTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3362)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAACCCGCCATCTGGTGGCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4935)	1	STAG1_+_136471182.23-P1P2	STAG1	P1P2	na	na	na	na	na	na	na	na	na
GAACTGCTGGGGAAGGCGAT (SEQ ID NO: 1790)	TATGAACCACTAAGGCGTccaccttgttgGAACTGCTGGGGAAGGCGATgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3363)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACATCGCCTTCCCCAGCAGTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4936)	1	STAG2_-_123095669.23-P1	STAG2	P1	na	na	na	na	na	na	na	na	na
GAGTCGGGCCTCGGAAGGGA (SEQ ID NO: 1791)	TATGAACCACTAAGGCGTccaccttgttgGAGTCGGGCCTCGGAAGGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3364)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCCTTCCGAGGCCCGACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4937)	1	STAG2_+_123095681.23-P1	STAG2	P1	na	na	na	na	na	na	na	na	na
GCGCGCTGGAGATGATGGGG (SEQ ID NO: 1792)	TATGAACCACTAAGGCGTccaccttgttgGCGCGCTGGAGATGATGGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3365)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCCATCATCTCCAGCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4938)	1	STAG2_+_123095617.23-P1	STAG2	P1	na	na	na	na	na	na	na	na	na
GCACCCTCGAATTCTACTCA (SEQ ID NO: 1793)	TATGAACCACTAAGGCGTccaccttgttgGCACCCTCGAATTCTACTCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3366)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGAGTAGAATTCGAGGGTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4939)	1	STAG2_+_123094566.23-P2	STAG2	P2	na	na	na	na	na	na	na	na	na
GGTCCCTGAGTAGAATTCGA (SEQ ID NO: 1794)	TATGAACCACTAAGGCGTccaccttgttgGGTCCCTGAGTAGAATTCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3367)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGAATTCTACTCAGGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4940)	1	STAG2_-_123094585.23-P2	STAG2	P2	na	na	na	na	na	na	na	na	na
GTCAGGGACGGAAGCTCCCG (SEQ ID NO: 1795)	TATGAACCACTAAGGCGTccaccttgttgGTCAGGGACGGAAGCTCCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3368)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGAGCTTCCGTCCCTGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4941)	1	STAG2_+_123094550.23-P2	STAG2	P2	na	na	na	na	na	na	na	na	na
GCCCGCCGCTGCCCTTCCGG (SEQ ID NO: 1796)	TATGAACCACTAAGGCGTccaccttgttgGCCCGCCGCTGCCCTTCCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3369)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGAAGGGCAGCGGCGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4942)	1	STAG3_+_99775559.23-P1P2	STAG3	P1P2	na	na	na	na	na	na	na	na	na
GGAATCTTTCGCCCCCGGAA (SEQ ID NO: 1797)	TATGAACCACTAAGGCGTccaccttgttgGGAATCTTTCGCCCCCGGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3370)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCCGGGGGCGAAAGATTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4943)	1	STAG3_-_99775570.23-P1P2	STAG3	P1P2	na	na	na	na	na	na	na	na	na
GGCCCCCGGAAGGGCAGCGG (SEQ ID NO: 1798)	TATGAACCACTAAGGCGTccaccttgttgGGCCCCCGGAAGGGCAGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3371)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCTGCCCTTCCGGGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4944)	1	STAG3_-_99775579.23-P1P2	STAG3	P1P2	na	na	na	na	na	na	na	na	na
GCGACCCGGGAGAAAGAAGT (SEQ ID NO: 1799)	TATGAACCACTAAGGCGTccaccttgttgGCGACCCGGGAGAAAGAAGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3372)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACTTCTTTCTCCCGGGTCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4945)	1	SUMO1_-_203102886.23-P1P2	SUMO1	P1P2	na	na	na	na	na	na	na	na	na
GCGGGCTGGGAGCCATCCAA (SEQ ID NO: 1800)	TATGAACCACTAAGGCGTccaccttgttgGCGGGCTGGGAGCCATCCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3373)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGGATGGCTCCCAGCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4946)	1	SUMO1_+_203103136.23-P1P2	SUMO1	P1P2	na	na	na	na	na	na	na	na	na
GGAAGTGACGCGAGGCGTAG (SEQ ID NO: 1801)	TATGAACCACTAAGGCGTccaccttgttgGGAAGTGACGCGAGGCGTAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3374)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTACGCCTCGCGTCACTTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4947)	1	SUMO1_+_203103297.23-P1P2	SUMO1	P1P2	na	na	na	na	na	na	na	na	na
GCGAGGCGCCAGCCAATGGG (SEQ ID NO: 1802)	TATGAACCACTAAGGCGTccaccttgttgGCGAGGCGCCAGCCAATGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3375)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCATTGGCTGGCGCCTCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4948)	1	SUMO2_+_73179116.23-P1P2	SUMO2	P1P2	na	na	na	na	na	na	na	na	na

GGCGGCGTGCGGCCC GTGG (SEQ ID NO: 1803)	TATGAACCACTAAGGCGTcaacctgtggGgcggagcggcgcgt	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCACGGG	SUMO2_-_7317880.42 3-P1P2	SUMO2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTTTCCCGGGCCCC CAG (SEQ ID NO: 1804)	TATGAACCACTAAGGCGTcaacctgtggGGCTTcaggggcccc	TCACCAGCAGAAGCCTTAGCTCTTAAACGTGGGGG	SUMO2_+_7 3778736.23-P1P2	SUMO2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTTGGGCCGGCCCACT TAG (SEQ ID NO: 1805)	TATGAACCACTAAGGCGTcaacctgtggGCTTCGGCGGGGCC	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCTAAGTC	SUMO3_-_46237725.2 3-P1P2	SUMO3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGCGCGCAACAGTTG CGG (SEQ ID NO: 1806)	TATGAACCACTAAGGCGTcaacctgtgCGCCGCCGCGCAACA	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCGCAACT	SUMO3_+_4 6237964.23-P1P2	SUMO3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTAGGGGCGGGGCTCT AGA (SEQ ID NO: 1807)	TATGAACCACTAAGGCGTcaacctgtgTTAggggcggggcgtcta gagtcaaggaGcaagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3386)	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTCTAGAGC	SUMO3_-_46237741.2 3-P1P2	SUMO3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGAGAATCACGGAGG TCC (SEQ ID NO: 1808)	TATGAACCACTAAGGCGTcaacctgtgGCGAGAATCACGG	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACGGAACCTC	SUPT16H_-_21852135.2 3-P1P2	SUPT16H	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGTCATTGGGAGAG TGC (SEQ ID NO: 1809)	TATGAACCACTAAGGCGTcaacctgtgCGGGGTCATTGGG	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACGACTCTCT	SUPT16H_+_21852155.2 3-P1P2	SUPT16H	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGGGAGAATGCGGGGAT TCC (SEQ ID NO: 1810)	TATGAACCACTAAGGCGTcaacctgtgGTGGGAGAATGCGG	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACGGAATCC	SUPT16H_+_21852157.2 3-P1P2	SUPT16H	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGCTCCGGGGGACCG AGG (SEQ ID NO: 1811)	TATGAACCACTAAGGCGTcaacctgtgGAGGGCTCCGGGGCGG	TCACCAGCAGAACTTTACGCAGCGTTAGCTCTTAAACCCTGGGT	SUZ12_+_30264157.23-P1P2	SUZ12	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGGGGGAACCGACG GGGA (SEQ ID NO: 1812)	TATGAAGCACTAAGGCGTcaacctgtgCCGGGGCGACCGA	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTCGCCCGT	SUZ12_+_30264152.23-P1P2	SUZ12	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGGGGGTAGCTGGC 93G (SEQ ID NO: 1813)	TATGAACCACTAAGGCGTcaacctgtgCCGGGGCTAGCTG	TCACCAGCAGAGCTTTAGCGAGCTTTAAACCCCGGCA	SUZ12_-_30264247.23-P1P2	SUZ12	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGCGGCTGGCAGTT CAA (SEQ ID NO: 1814)	TATGAACCACTAAGGCGTcaacctgtgGCGGGCGGCTGGC	TCACCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACTTGAACTG	SYCP1_+_115397462.23-P1P2	SYCP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGAGCGGCCCGGAGCG CGT (SEQ ID NO: 1815)	TATGAAGCACTAAGGCGTcaacctgtgCGTGAGCGGGCCCGG	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACACGGGCT	SYCP1_+_115397487.23-P1P2	SYCP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCTACATAGCGTCTAT GA (SEQ ID NO: 1816)	TATGAACCACTAAGGCGTcaacctgtgGTGCTACATAGCGT	TCACCAGCAGAGCTTTAGCGAGCGTTAAAGTCATAGACT	SYCP1_+_115397571.23-P1P2	SYCP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGAGCGCTCGGGCGCAT AGG (SEQ ID NO: 1817)	TATGAACCACTAAGGCGTcaacctgtgAGCGAGCGCTCGGG	TCACCAGCAGAGCTTTAGCGAGCGTTAGTGGTTCATA	SYCP2_+_58538681.23-P1P2	SYCP2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGCTCCACCGACTACT ACG (SEQ ID NO: 1818)	TATGAAGCACTAAGGCGTcaacctgtgCGCGGCTCCACCGA	TCAGCAGCAGAGCTTTAGCGAGCTTAGCTCTTAAACCCTACTGT	SYCP2_+_58539539.23-P1P2	SYCP2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGCGGTCCGCGGTAG TGT (SEQ ID NO: 1819)	TATGAACCACTAAGGCGTcaacctgtgGGAAGCGGTCCGCG	CGGACGCGCTCCCCAACAAGCGTTAGCTCTTAAACACACTAGG	SYCP2_-_58539882.23-P1P2	SYCP2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCAGAAGCGCCGCTTCAC CTG (SEQ ID NO: 1820)	TATGAAGCACTAAG3GCGTcaacctgtgGCACAAGCGGCGT	TCACCTgttaagagtaagngtGTAAAGTCTGCTGTGA (SEQ ID NO: 3389)	TCAGCAGCAGAGCTTACGCAGCTTAGGTCTTAAACCAGGTGA	1	SYCP3_-_102133215 23-P1P2	SYCP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGCTTGTGTGAGGAG TTTG (SEQ ID NO: 1821)	TATGAACCACTAAGGCGTcaacctgtgGCGCGCTTGTGAGA	TGAGTTTgttaagagtaagngtaagngtGTAAAGTCTGTGTGA (SEQ ID NO: 3390)	TCAACGAGCAGAGCTTACGCAGCTTAGGTCTTAAACCAACTCCT	1	SYCP3_+_102133182 23-P1P2	SYCP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CGCCGGAGAACCGCGCCAC CTC (SEQ ID NO: 1822)	TATGAACCACTAAGGCGTcaacctgtgGGCGAGAAGGGGG	CCACCTCgttaagagtaagngtGTAAAGTCTGTGGTGA (SEQ ID NO: 3391)	TCAACGAGCAGAGCTTACGCAGCTTAGGTCTTAAACGAGTGTG	1	SYCP3_+_102133259 23-P1P2	SYCP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGTAGCCGACATCAGCT AA (SEQ ID NO: 1823)	TATGAAGCACTAAG3GCGTcaacctgtgGACCTACCGCGACAT	GACTAAgttaagagtaagngtaagngtGGTAAAGTCTGTGGTGA (SEQ ID NO: 3392)	TCAGCAGCAGAGCTTACGCAGCTTAGGTCTTAAACTTAGTCAT	1	TAF15_+_34155664 23-P1P2	TAF15	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGCGCTCAGTAGTAG CTC (SEQ ID NO: 1824)	TATGAACCACTAAGGCGTcaacctgtgGGGCGCCTCAGTT	ACTCAGCCTCgttaagagtaagngtaagngtGGTAAAGTCTGTGGTGA (SEQ ID NO: 3393)	TCAGCAGCAGAGCTTACGCAGCTTAGGTCTTAAACGAGTGTGT	1	TAF15_-_34139469 23-P1P2	TAF15	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGGCGGGAGCTGTA CTG (SEQ ID NO: 1825)	TATGAACCACTAAGGCGTcaacctgtgGGGCGGGAGGA	TCACCTgttaagagtaagngtaagngtGGTAAAGTCTGTGGTGA (SEQ ID NO: 3394)	TCACCAGCAGAGCTTACGCAGCTTAGGTCTTAAACCAGTACA	1	TAF15_+_34136463 23-P1P2	TAF15	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACTGCAGATCGCTGGT GAG (SEQ ID NO: 1826)	TATGAAGCACTAAG3GCGTcaacctgtgCAGTGCACATCGC	TGCTGAGgttaagagtaagngtaagngtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3395)	TCAGCAGCAGAGCTTACGCAGCTTAGGTCTTAAACTGCACGAG	1	TCEA1_-_54934845 23-P1P2	TCEA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTTCGGTTCGTTAAGGAA 393 (SEQ ID NO: 1827)	TATGAACCACTAAGGCGTcaacctgtgGGCTTTCGTTAAA	CCCgaggtttaagagtaagngtaagngtCGTAAAGTCTGGTTGA (SEQ ID NO: 3401)	TCACCAGCAGAGCTTACGCAGCTTAGGTCTTAAACCCTTCCTT	1	TCEA1_+_54934842 23-P1P2	TCEA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTACGAAGCGAAGCCCG CGG (SEQ ID NO: 1828)	TATGAAGCACTAAG3GCGTcaacctgtgGTTACGAAGCGAAG	CCCgaggtttaagagtaagngtaagngtCGTAAAGTCTGGTTGA (SEQ ID NO: 3401)	TCAGCAGCAGAGCTTACGCAGCTTAGGTCTTAAACCGCGCGG	1	TCEA1_-_54934842 23-P1P2	TCEA1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGACGCTGGGCCCGAT GGC (SEQ ID NO: 1829)	TATGAAGCACTAAG3GCGTcaacctgtgCGACGCCTGGGCC	TCAGCAGCAGAGCTTACGCAGCTTAGGTCTTAAACGCCATCG	1	TDG_+_1043569525 23-P1P2	TDG	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTGGCAGTACCCCAGA CAG (SEQ ID NO: 1830)	TATGAACCACTAAGGCGTcaacctgtgGGCTGGCAGTACC	GGTACTGCCCAGGCCCAAACGTCTGCTGGTGA (SEQ ID NO: 3403)	TCAGCAGCAGAGCTTTACGCAGCTTAGGTCTTAAACCTGTCTG	1	TDG_+_1043599778 23-P1P2	TDG	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCTGGGGTACTGCCCAGC CAT (SEQ ID NO: 1831)	TATGAACCACTAAGGCGTcaacctgtgGTCTGGGGTACTGC	CAGCCCAgttaagagtaagngtaagngtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3404)	TCAGCAGCAGAGCTTTACGCAGCTTAGGTCTTAAACATGGGCTG	1	TDG_-_104359705 23-P1P2	TDG	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGAATCCGAGCGCAAG CGT (SEQ ID NO: 1832)	TATGAAGCACTAAG3GCGTcaacctgtgCGCGGGAATCCGAG	CACTCGGATCCCGCGCTTGCTGGTGA (SEQ ID NO: 3405)	TCAGCAGCAGAGCTTACGCAGCTTAGGTCTTAAACGCGGTTG	1	TDP1_-_90422298 23-P1P2	TDP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGAAGGAGCGCGGGAT CCG (SEQ ID NO: 1833)	TATGAACCACTAAGGCGTcaacctgtgGGCCGAAGGGGGTG	CGGATCCgttaagagtaagngtaagngtCGTAAAGTCTGTGGTGA (SEQ ID NO: 3406)	TCACCAGCAGAGCTTTACGCAGCTTAGGTCTTAAACGCGGGATCC	1	TDP1_-_90422288 23-P1P2	TDP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGCCCAACAGCGTTAC CTG (SEQ ID NO: 1834)	TATGAACCACTAAGGCGTcaacctgtgGGCCGCCCAACAG	TGACCTGgttaagagtaagngtaagngtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3407)	TGACCAGCAGAGCTTTACGCAGCTTAGGTCTTAAACCAGGTAC	1	TDP1_+_90422383 23-P1P2	TDP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTCCAGCCTAAGGAG TAA (SEQ ID NO: 1835)	TATGAAGCACTAAG3GCGTcaacctgtgGGCTCCAGCGAGCC	TAGGTAAgttaagagtaagngtaagngtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3408)	TCAGCAGCAGAGCTTTACGCAGCTTAGGTCTTAAACTTAGCTAG	1	TDP2_-_24869895 23-P1P2	TDP2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCGGGCAGGAAGAGTGG AGT (SEQ ID NO: 1836)	TATGAACCACTAAGGCGTcaacctgtgGGGCGGCAGCGAAGC	ATGCAGTgttaagagtaagngtaagngtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3409)	TCACCAGCAGAGCTTTACGCAGCTTAGGTCTTAAACGCTGCCATC	1	TDP2_+_24567080 23-P1P2	TDP2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTAAAGGGGCGGTGCAGAGG (SEQ ID NO: 1837)	TATGAACCACTAAGGCGTccaccttgttgGTAAAGGGGCGGTGCAGAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3410)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTCTGCACCGCCCCTTTACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4983)	1	TDP2_+_24667097.23-P1P2	TDP2	P1P2	na	na	na	na	na	na	na	na	na
GGCCCCGAGCACGAGTTCCT (SEQ ID NO: 1838)	TATGAACCACTAAGGCGTccaccttgttgGGCCCCGAGCACGAGTTCCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3411)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGGAACTCGTGCTCGGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4984)	1	TELO2_+_1543463.23-P1P2	TELO2	P1P2	na	na	na	na	na	na	na	na	na
GGCCTCTCAGCGCGAATCTG (SEQ ID NO: 1839)	TATGAACCACTAAGGCGTccaccttgttgGGCCTCTCAGCGCGAATCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3412)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGATTCGCGCTGAGAGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4985)	1	TELO2_+_1543388.23-P1P2	TELO2	P1P2	na	na	na	na	na	na	na	na	na
GGTCCGGGCGAGGCTCCCGC (SEQ ID NO: 1840)	TATGAACCACTAAGGCGTccaccttgttgGGTCCGGGCGAGGCTCCCGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3413)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGGGAGCCTCGCCCGGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4986)	1	TELO2_+_1543442.23-P1P2	TELO2	P1P2	na	na	na	na	na	na	na	na	na
GAGCGAGCCATTTAACATGG (SEQ ID NO: 1841)	TATGAACCACTAAGGCGTccaccttgttgGAGCGAGCCATTTAACATGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3414)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCATGTTAAATGGCTCGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4987)	1	TERF1_-_73921127.23-P1P2	TERF1	P1P2	na	na	na	na	na	na	na	na	na
GCGTTGCTCGGCGCCTGAAG (SEQ ID NO: 1842)	TATGAACCACTAAGGCGTccaccttgttgGCGTTGCTCGGCGCCTGAAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3415)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTTCAGGCGCCGAGCAACGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4988)	1	TERF1_-_73921096.23-P1P2	TERF1	P1P2	na	na	na	na	na	na	na	na	na
GGCGGAGGATGTTTCCTCAG (SEQ ID NO: 1843)	TATGAACCACTAAGGCGTccaccttgttgGGCGGAGGATGTTTCCTCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3416)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGAGGAAACATCCTCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4989)	1	TERF1_-_73921145.23-P1P2	TERF1	P1P2	na	na	na	na	na	na	na	na	na
GAGTAGCGACGGCAGCGGGC (SEQ ID NO: 1844)	TATGAACCACTAAGGCGTccaccttgttgGAGTAGCGACGGCAGCGGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3417)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCCGCTGCCGTCGCTACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4990)	1	TERF2_+_69419708.23-P1P2	TERF2	P1P2	na	na	na	na	na	na	na	na	na
GATGGCCGCGGGAGCCGGGA (SEQ ID NO: 1845)	TATGAACCACTAAGGCGTccaccttgttgGATGGCCGCGGGAGCCGGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3418)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCCGGCTCCCGCGGCCATCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4991)	1	TERF2_+_69419852.23-P1P2	TERF2	P1P2	na	na	na	na	na	na	na	na	na
GGAGGAGGCGGGAGTAGCGA (SEQ ID NO: 1846)	TATGAACCACTAAGGCGTccaccttgttgGGAGGAGGCGGGAGTAGCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3419)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCTACTCCCGCCTCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4992)	1	TERF2_+_69419719.23-P1P2	TERF2	P1P2	na	na	na	na	na	na	na	na	na
GAACTGAAGAGCGGCGGCGT (SEQ ID NO: 1847)	TATGAACCACTAAGGCGTccaccttgttgGAACTGAAGAGCGGCGGCGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3420)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACGCCGCCGCTCTTCAGTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4993)	1	THOC1_-_268012.23-P1P2	THOC1	P1P2	na	na	na	na	na	na	na	na	na
GCGGGCAAACTGAAGAGCGG (SEQ ID NO: 1848)	TATGAACCACTAAGGCGTccaccttgttgGCGGGCAAACTGAAGAGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3421)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCTCTTCAGTTTGCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4994)	1	THOC1_-_268006.23-P1P2	THOC1	P1P2	na	na	na	na	na	na	na	na	na
GGGCGTCGGAGACATCTTCT (SEQ ID NO: 1849)	TATGAACCACTAAGGCGTccaccttgttgGGGCGTCGGAGACATCTTCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3422)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGAAGATGTCTCCGACGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4995)	1	THOC1_-_268026.23-P1P2	THOC1	P1P2	na	na	na	na	na	na	na	na	na
GCTGCACTCTCACCTGTAGT (SEQ ID NO: 1850)	TATGAACCACTAAGGCGTccaccttgttgGCTGCACTCTCACCTGTAGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3423)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACTACAGGTGAGAGTGCAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4996)	1	THOC2_+_122866726.23-P1P2	THOC2	P1P2	na	na	na	na	na	na	na	na	na
GGAGTTTCGAACTCGGGGAG (SEQ ID NO: 1851)	TATGAACCACTAAGGCGTccaccttgttgGGAGTTTCGAACTCGGGGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3424)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCCCGAGTTCGAAACTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4997)	1	THOC2_+_122866641.23-P1P2	THOC2	P1P2	na	na	na	na	na	na	na	na	na
GGGCTTCTGCTACTAGTGAG (SEQ ID NO: 1852)	TATGAACCACTAAGGCGTccaccttgttgGGGCTTCTGCTACTAGTGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3425)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCACTAGTAGCAGAAGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4998)	1	THOC2_+_122866875.23-P1P2	THOC2	P1P2	na	na	na	na	na	na	na	na	na
GAGCAAGACGCGCGAGTTCC (SEQ ID NO: 1853)	TATGAACCACTAAGGCGTccaccttgttgGAGCAAGACGCGCGAGTTCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3426)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGAACTCGCGCGTCTTGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 4999)	1	THOC3_+_175395054.23-P1P2	THOC3	P1P2	na	na	na	na	na	na	na	na	na

GCAACTTTCGCGCTGTGC (SEQ ID NO: 1854)	TATGAAGCACTAAGGCGTcaacttgtgGCAACGTTGGCGCT (GTCGCGCCaattaagagcaagtgtCTAAAGTCTGCTCGTGA (SEQ ID NO: 3427)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAAGCGGCGAC AGCGGCCAAAGGTGGCAACAAGGTGGAGCGCCTTAGTCGTTCATA (SEQ ID NO: 5002)	TH0C3_-_17539059 23-P1P2	TH0C3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCACGGCGCCATCGAG CCG (SEQ ID NO: 1855)	TATGAACCACTAAGGCGTcaacttgtgCCACGGCGCCAT (CGCGCCCGgttaagagctaagtgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3428)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCACTACCG ATCGGCCCGGTTGGCCAAGAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5001)	TH0C3_-_17539059 23-P1P2	TH0C3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGCGCAGCCGCGGTA GTG (SEQ ID NO: 1856)	TATGAACCACTAAGGCGTcaacttgtgCCGCGGCGCTGC (GCGGTCGCCGCGGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3429)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCACTACCG GCGGTCGCCGCGGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5002)	TH0C5_-_29946540 2 3-P1P2	TH0C5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCACAGCTTCCAACGAA CC (SEQ ID NO: 1857)	TATGAACCACTAAGGCGTcaacttgtgCCCTTCACAGTTCCA (GAGGTCTGTGACGGCCCAACAAGGTTCGACGCCTTAGTCGTTCATA (SEQ ID NO: 5003)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCGTCTTG GAGGTCTGTGACGGCCCAACAAGGTTCGACGCCTTAGTCGTTCATA (SEQ ID NO: 5003)	TH0C5_-_29946579 2 3-P1P2	TH0C5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCAGTAAGCCACTGC CGG (SEQ ID NO: 1858)	TATGAACCACTAAGGCGTcaacttgtgGGGCAGTTAAGCCA (TGTGCTGAgttaagagctaagtgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3431)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCCGGCAG TGTGCTTACTCCCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5004)	TH0C5_-_29945705 2 3-P1P2	TH0C5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGACTTACGTGCCTA GAG (SEQ ID NO: 1859)	TATGAACCACTAAGGCGTcaacttgtgGAGGCCTTACGTG (CCTAGAgttaagagctaagtgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3432)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCTCTAGG CACGTTAAGGCCTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5005)	TH0C6_+_3 074256 23-P1P2	TH0C6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCGGGCGCGGAAGAC CGG (SEQ ID NO: 1860)	TATGAAGCACTAAGGCGTcaacttgtgCGCGCGCGCGGA (AGACGGGgttaagagctaagtgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3433)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCCGGTCTT CCCGCGCCCGGCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5006)	TH0C6_-_3974073 23-P1P2	TH0C6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACGCCGCGCAGAAGC TCG (SEQ ID NO: 1861)	TATGAACCACTAAGGCGTcaacttgtgGGCGCCGCGCAG (AACCTTCgttaagagctaagtgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3434)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACGAGGTT CTTGCGCGGTTGCGCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5007)	TH0C6_+_3 074038 23-P1P2	TH0C6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCATGGGAGCCGTGAC TGA (SEQ ID NO: 1862)	TATGAACCACTAAGGCGTcaacttgtgGCCATGGGAGCCG (TGAGTCGgttaagagctaagtgtCGTAAAGTCTGCTCGTGA (SEQ ID NO: 3435)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACTCAGTGAC GGGTCCCATGGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5008)	TH0C7_+_6 3849446 23-P1P2	TH0C7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCAACCTCGAGCGCGC GGT (SEQ ID NO: 1863)	TATGAAGCACTAAGGCGTcaacttgtgCGTAAAGTCTGCTGGTGA (SEQ ID gtgttaagagctaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3436)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCACCGCGT CCTCAAGCTTGCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5009)	TH0C7_-_53849518 2 3-P1P2	TH0C7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACGGCGGCGCCAAGCTG AGG (SEQ ID NO: 1864)	TATGAACCACTAAGGCGTcaacttgtgCGCGCGGCGGGA (GGTTAAGgtgtaagctgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3437)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCGTCAGGT TGCCGCGCCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5010)	TH0C7_-_53849512 2 3-P1P2	TH0C7	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACGGAGCCCGCTGAG AAG (SEQ ID NO: 1865)	TATGAACCACTAAGGCGTcaacttgtgGAACGGAGCCCGC (TCAGAAGgttaagagctaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3438)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCTTCTGAG CGGCGCTCCGTTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5011)	TH1RAP3_+_36890245 23-P1P2	TH1RAP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCTGCGCGCGCTGTG GGG (SEQ ID NO: 1866)	TATGAAGCACTAAGGCGTcaacttgtgCGCGCTGCGCGC (TGTGGGGgttaagagctaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3439)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCCCCAGCA GGCGGGGCCCAGCGGCCAAACAAGGTCGACGCCTTAGTGGTTCAT A (SEQ ID NO: 5012)	TH1RAP3_-_38690342 2 3-P1P2	TH1RAP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTGCGACGCTGCCATGT CTG (SEQ ID NO: 1867)	TATGAACCACTAAGGCGTcaacttgtgGCTGCGACGCTGC (ATCTCTGgttaagagctaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3440)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACGCAGAGAT CGCGCTGCGACGCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5013)	TH1RAP3_-_38690366 2 3-P1P2	TH1RAP3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACGCGCGGCGGGGAGCA CTA (SEQ ID NO: 1868)	TATGAACCACTAAGGCGTcaacttgtgGAACGCGCGGCGG (GAGACTAgttaagagctaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3441)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACTAGTCTCC CGCGCGCGGGTTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5014)	TIMELESS_-_58942828 2 3-P1P2	TIMELESS	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGACGCCCGCGAGGAGC CAC (SEQ ID NO: 1869)	TATGAAGCACTAAGGCGTcaacttgtgCGACGACGCCCGG (GAGCCACgttaagagctaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3442)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACGTTCAG CTCGCGGGCTTCTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5015)	TIMELESS_-_58843133 2 3-P1P2	TIMELESS	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCTGACGACCCGCGGA ACG (SEQ ID NO: 1870)	TATGAACCACTAAGGCGTcaacttgtgCCGCGCTGACGA (CCGCAACgttaagagctaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3443)	TCACCAGCAGAGACTTTAGCGCAGCTTAGCTCTTAAACCGCTTCCG GGCGTCTCACCCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5016)	TIMELESS_-_58942914 2 3-P1P2	TIMELESS	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCAACTCAACGCACAAAA CAC (SEQ ID NO: 1971)	TATGAAACACATAAGGCGTcaacctgtgGCAACGTCAGGCAG AAAACACAgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3444)	TCAACAACACAGAGTTTACGGCAAGGTTAGCTCTTAAACGTTGTTTTC 1 TGCGTGAG391TGCCAACAAG391TGGACCGCCTTAGTCGCTTCATA (SEQ ID NO: 5017)	TTPN_-_86648694 2 3-P1P2	TTPN	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCGATAGTCGTGGAACT CCT (SEQ ID NO: 1972)	TATGAACCACTAAGGCGTcaacctgtgGCGATAGTCTGGTGA AACTCCTgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3445)	TCAACAAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACATGAGGTT 1 CCCGAGTATCGGCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5018)	TTPN_-_86645067 2 3-P1P2	TTPN	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAACTCCTCCGGCCGA GAA (SEQ ID NO: 1973)	TATGAACCACTAAGGCGTcaacctgtgGGAAACTCCTGGTGA GCGGAGAAgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3446)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACTTCTCGGCC 1 CGAGGAGTTCCCCAAGAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5019)	TTPN_-_86645067 2 3-P1P2	TTPN	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTCGGATCCCCCGCCGG ATC (SEQ ID NO: 1974)	TATGAACACACTAAGGCGTcaacctgtgCGTCCGGATCCGCCG CGGCGATCCTgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3447)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACGGATCCCCG 1 CGGCGATCCCGAGCCAACAAG391TGGACGGCTTAGTGGTTCAT A (SEQ ID NO: 5020)	TONSL_-_14566969 19 23-P1P2	TONSL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTTCGCCGTAAGTTCC TTC (SEQ ID NO: 1975)	TATGAACCACTAAGGCGTcaacctgtgGGGTTCGCCGTAA GTTCCTCgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3448)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACTAAGTTCC 1 TTACGGCGAAGCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5021)	TONSL_+_1 4566978 23-P1P2	TONSL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGGGTCCCCGAGGAAC TTA (SEQ ID NO: 1976)	TATGAACCACTAAGGCGTcaacctgtgGGGGGTCCCCGAG GAACTTAgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3449)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACTAAGTTCC 1 TCGGGGGACCCCCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5022)	TONSL_-_14566977 23-P1P2	TONSL	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGGTGGCTTGACTTGCC CGG (SEQ ID NO: 1977)	TATGAAGCACTAAGGCGTcaacctgtgCGACGCTGCCTGAC CGGCGGAgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3450)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACCCGGCGCA 1 GTCAGGCAAGCGTCCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5023)	TOP1_+_386 57511 23-P1P2	TOP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGGTTGGTCCCCACTCA TGT (SEQ ID NO: 1978)	TATGAACCACTAAGGCGTcaacctgtgGAGGTTGGTCCCCA CTCATGTgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3451)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACCACATGAGTT 1 GGGGAGCAACCTCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5024)	TOP1_+_386 57702 23-P1P2	TOP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCAGGGCAGGCGTCGGCG CCG (SEQ ID NO: 1979)	TATGAACCACTAAGGCGTcaacctgtgGTCAGGGCAGGCGTC GGCGGCGgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3452)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACCGGCGGC 1 GACGGCTGCCTGACCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5025)	TOP1_-_39657543 2 3-P1P2	TOP1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGGCGCTCTGTCGCAG AAG (SEQ ID NO: 1980)	TATGAAGCACTAAGGCGTcaacctgtgGACGCGCTCTGTCG CAGAAGgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3453)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACCTTCTCCGA 2 CAGACGCGCGCTCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5026)	TOP2A_+_38 5741211 23-P1P2	TOP2A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAGAGAAGCGGGCTTGGT CGG (SEQ ID NO: 1981)	TATGAACCACTAAGGCGTcaacctgtgGGGAGAAGAGGGCT TGGTCGGgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3454)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTTAAACGCGGACGA 2 AGCCGCTTCTCCCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5027)	TOP2A_+_38 5741109 23-P1P2	TOP2A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGGCGGCCAGAGAGATG CCC (SEQ ID NO: 1982)	TATGAACCACTAAGGCGTcaacctgtgGGGCGGCCAGAGAG ATGCCCgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3455)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACGGGGCAATC 2 TCTCTGGCGCGCCCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5028)	TOP2A_-_385739792 3-P1P2	TOP2A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GACGGGAGGTGACCGTGG CGG (SEQ ID NO: 1983)	TATGAAGCACTAAGGCGTcaacctgtgGAGGGAGGTGACC GTGGCGGgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3456)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACCCGGCGAC 1 GGTCACCTCCCCTGCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5029)	TOP2B_+_25 7763650 23-P1P2	TOP2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGTGTTTATGCGCGGGAGA AGA (SEQ ID NO: 1984)	TATGAACCACTAAGGCGTcaacctgtgGGGTGTTTATGCGC GGGAGAAgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3457)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAAGTGTTGTCC 1 GGCATAACACCCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5030)	TOP2B_+_25 7763661 23-P1P2	TOP2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTTGCACACGCGGGGAG AGA (SEQ ID NO: 1985)	TATGAACCACTAAGGCGTcaacctgtgGTTTGCACACGCGG GGAGAAgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3458)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACTCTGTGCC 1 CGGTGTGGCAAACCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5031)	TOP2B_-_25705422 2 3-P1P2	TOP2B	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAGTTCCACGCCAACGAT AC (SEQ ID NO: 1986)	TATGAACCACTAAGGCGTcaacctgtgCGAAGTTCCACGCCA ACGATACgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3459)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACGGTGGACA 1 GCGTGGGAACTTCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5032)	TOP3A_+_18 2192284 23-P1P2	TOP3A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
CAACTGGAATTCGGTTCGTG AGG (SEQ ID NO: 1987)	TATGAACCACTAAGGCGTcaacctgtgGAGTTCGGTTCGTG AGGgttaagagtaagtaagtgtCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3460)	TCAACGAGCAGAGTTTACGGCAAGGTTAGCTCTTAAACCGTCAGGA 1 CCGAATTCCACGTGCCAACAAG391TGGACGGCTTAGTGGTTCATA (SEQ ID NO: 5033)	TOP3A_+_18 2192267 23-P1P2	TOP3A	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GTCAGTGACTGAAGCTCCAA (SEQ ID NO: 1888)	TATGAACCACTAAGGCGTccaccttgttgGTCAGTGACTGAAGCTCCAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3461)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTGGAGCTTCAGTCACTGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5034)	1	TOP3A_+_18218224.23-P1P2	TOP3A	P1P2	na	na	na	na	na	na	na	na	na
GCGCACCGCGGATCCAGCTC (SEQ ID NO: 1889)	TATGAACCACTAAGGCGTccaccttgttgGCGCACCGCGGATCCAGCTCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3462)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGAGCTGGATCCGCGGTGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5035)	1	TOP3B_-_22337154.23-P1P2	TOP3B	P1P2	na	na	na	na	na	na	na	na	na
GCGGCTGGACCCGGAAATAG (SEQ ID NO: 1890)	TATGAACCACTAAGGCGTccaccttgttgGCGGCTGGACCCGGAAATAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3463)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTATTTCCGGGTCCAGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5036)	1	TOP3B_+_22337183.23-P1P2	TOP3B	P1P2	na	na	na	na	na	na	na	na	na
GGGTAGCCGCCCCGGGAACA (SEQ ID NO: 1891)	TATGAACCACTAAGGCGTccaccttgttgGGGTAGCCGCCCCGGGAACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3464)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTTCCCGGGGCGGCTACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5037)	1	TOP3B_+_22337156.23-P1P2	TOP3B	P1P2	na	na	na	na	na	na	na	na	na
GGGCATCGCGAACCCCAGTT (SEQ ID NO: 1892)	TATGAACCACTAAGGCGTccaccttgttgGGGCATCGCGAACCCCAGTTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3465)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAACTGGGGTTCGCGATGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5038)	1	TOPBP1_-_133380571.23-P1P2	TOPBP1	P1P2	na	na	na	na	na	na	na	na	na
GGGCTCCAACGAGGTAAGTG (SEQ ID NO: 1893)	TATGAACCACTAAGGCGTccaccttgttgGGGCTCCAACGAGGTAAGTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3466)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCACTTACCTCGTTGGAGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5039)	1	TOPBP1_+_133380602.23-P1P2	TOPBP1	P1P2	na	na	na	na	na	na	na	na	na
GTCTCCGGGCGGCGAGTCGG (SEQ ID NO: 1894)	TATGAACCACTAAGGCGTccaccttgttgGTCTCCGGGCGGCGAGTCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3467)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGACTCGCCGCCCGGAGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5040)	1	TOPBP1_-_133380651.23-P1P2	TOPBP1	P1P2	na	na	na	na	na	na	na	na	na
GAAGAGGCGGAGTCGAGGTG (SEQ ID NO: 1895)	TATGAACCACTAAGGCGTccaccttgttgGAAGAGGCGGAGTCGAGGTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3468)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCACCTCGACTCCGCCTCTTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5041)	1	TOPORS_+_32552597.23-P1P2	TOPORS	P1P2	na	na	na	na	na	na	na	na	na
GCGCCTGGGAGGTTACTGTA (SEQ ID NO: 1896)	TATGAACCACTAAGGCGTccaccttgttgGCGCCTGGGAGGTTACTGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3469)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACAGTAACCTCCCAGGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5042)	1	TOPORS_-_32552385.23-P1P2	TOPORS	P1P2	na	na	na	na	na	na	na	na	na
GCGTATCCTCCAGGAAGAGG (SEQ ID NO: 1897)	TATGAACCACTAAGGCGTccaccttgttgGCGTATCCTCCAGGAAGAGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3470)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCTCTTCCTGGAGGATACGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5043)	1	TOPORS_+_32552610.23-P1P2	TOPORS	P1P2	na	na	na	na	na	na	na	na	na
GAGTCTAGAGCCACCGTCCA (SEQ ID NO: 1898)	TATGAACCACTAAGGCGTccaccttgttgGAGTCTAGAGCCACCGTCCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3471)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGACGGTGGCTCTAGACTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5044)	1	TP53_+_7590780.23-P1P2	TP53	P1P2	na	na	na	na	na	na	na	na	na
GCAGGGAGCAGGTAGCTGCT (SEQ ID NO: 1899)	TATGAACCACTAAGGCGTccaccttgttgGCAGGGAGCAGGTAGCTGCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3472)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGCAGCTACCTGCTCCCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5045)	1	TP53_+_7590763.23-P1P2	TP53	P1P2	na	na	na	na	na	na	na	na	na
GGAGCCACCGTCCAGGGAGC (SEQ ID NO: 1900)	TATGAACCACTAAGGCGTccaccttgttgGGAGCCACCGTCCAGGGAGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3473)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCTCCCTGGACGGTGGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5046)	1	TP53_+_7590774.23-P1P2	TP53	P1P2	na	na	na	na	na	na	na	na	na
GACCTAGGGATCGATCTGGA (SEQ ID NO: 1901)	TATGAACCACTAAGGCGTccaccttgttgGACCTAGGGATCGATCTGGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3474)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCCAGATCGATCCCTAGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5047)	1	TP53BP1_+_43785295.23-P1	TP53BP1	P1	na	na	na	na	na	na	na	na	na
GGGCGACCTAGGGATCGATC (SEQ ID NO: 1902)	TATGAACCACTAAGGCGTccaccttgttgGGGCGACCTAGGGATCGATCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3475)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGATCGATCCCTAGGTCGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5048)	1	TP53BP1_+_43785299.23-P1	TP53BP1	P1	na	na	na	na	na	na	na	na	na
GTGGCGACAGCGGCGACCTA (SEQ ID NO: 1903)	TATGAACCACTAAGGCGTccaccttgttgGTGGCGACAGCGGCGACCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3476)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGGTCGCCGCTGTCGCCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5049)	1	TP53BP1_+_43785309.23-P1	TP53BP1	P1	na	na	na	na	na	na	na	na	na
GAGCCTTCCAGAGCACTCCC (SEQ ID NO: 1904)	TATGAACCACTAAGGCGTccaccttgttgGAGCCTTCCAGAGCACTCCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3477)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGAGTGCTCTGGAAGGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5050)	1	TP53BP1_-_43802635.23-P2	TP53BP1	P2	na	na	na	na	na	na	na	na	na

GAGGCTGCGCAGTATTG GAG (SEQ ID NO: 1905)	TATGAACCACTAAGGCGTccaccttgttgGAGGCTGCGCAGT ATTGgaggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3478)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCAATA CTGCGCAGCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5051)	1	TP53BP1_+_43802597.23-P2	TP53BP1	P2	na	na	na	na	na	na	na	na	na
GTGCCCGAGGTCTCCAT CCT (SEQ ID NO: 1906)	TATGAACCACTAAGGCGTccaccttgttgGTGCCCGAGGTCT CCATCCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3479)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGGATGG AGACCTCGGGCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5052)	1	TP53BP1_+_43802658.23-P2	TP53BP1	P2	na	na	na	na	na	na	na	na	na
GCCGGCCGGGCCCCCTG ACG (SEQ ID NO: 1907)	TATGAACCACTAAGGCGTccaccttgttgGccggccgggcccccTG ACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3480)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTCAGG GGGCCCGGCCGGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 5053)	1	TP53BP2_-_224033594.23-P1P2	TP53BP2	P1P2	na	na	na	na	na	na	na	na	na
GGGATTAGTTGGTTTCGG AG (SEQ ID NO: 1908)	TATGAACCACTAAGGCGTccaccttgttgGGGATTAGTTGGTT TCGGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3481)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCGAAA CCAACTAATCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5054)	1	TP53BP2_+_224033636.23-P1P2	TP53BP2	P1P2	na	na	na	na	na	na	na	na	na
GGTCTCTTCGACGCTCGT GA (SEQ ID NO: 1909)	TATGAACCACTAAGGCGTccaccttgttgGGTCTCTTCGACG CTCGTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3482)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCACGAG CGTCGAAGAGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5055)	1	TP53BP2_-_224033620.23-P1P2	TP53BP2	P1P2	na	na	na	na	na	na	na	na	na
GGCCCTGGGCCGGAGAA CCC (SEQ ID NO: 1910)	TATGAACCACTAAGGCGTccaccttgttgGGCCCTGGGCCGG AGAACCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3483)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGTTCT CCGGCCCAGGGCCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 5056)	1	TP53I11_+_44971590.23-P1P2	TP53I11	P1P2	na	na	na	na	na	na	na	na	na
GGCTTGTGCGCGGAACC GAG (SEQ ID NO: 1911)	TATGAACCACTAAGGCGTccaccttgttgGGCTTGTGCGCGG AACCGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3484)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCGGTTC CGCGCACAAGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5057)	1	TP53I11_+_44972454.23-P1P2	TP53I11	P1P2	na	na	na	na	na	na	na	na	na
GGGTAAGCCCCGGCCGC GCA (SEQ ID NO: 1912)	TATGAACCACTAAGGCGTccaccttgttgGGGTAAGCCCCGG CCGCGCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3485)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGCGCGG CCGGGGCTTACCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5058)	1	TP53I11_+_44972406.23-P1P2	TP53I11	P1P2	na	na	na	na	na	na	na	na	na
GCAGGCAGGCAGTAGTG ACT (SEQ ID NO: 1913)	TATGAACCACTAAGGCGTccaccttgttgGcaggcaggcagTAGT GACTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3486)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGTCACTA CTGCCTGCCTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5059)	1	TREX1_+_48507673.23-P1P2	TREX1	P1P2	na	na	na	na	na	na	na	na	na
GCTCACCGTAGCAGGCA GGC (SEQ ID NO: 1914)	TATGAACCACTAAGGCGTccaccttgttgGCTCACCGTAgcagg caggcgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3487)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCTGCC TGCTACGGTGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5060)	1	TREX1_+_48507691.23-P1P2	TREX1	P1P2	na	na	na	na	na	na	na	na	na
GTCGGCTGCACACTCTC CCG (SEQ ID NO: 1915)	TATGAACCACTAAGGCGTccaccttgttgGTCGGCTGCACAC TCTCCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3488)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGAGA GTGTGCAGCCGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5061)	1	TREX1_+_48507655.23-P1P2	TREX1	P1P2	na	na	na	na	na	na	na	na	na
GCGGCTCGGAGCGCGAA ACA (SEQ ID NO: 1916)	TATGAACCACTAAGGCGTccaccttgttgGCGGCTCGGAGCG CGAAACAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3489)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGTTTCGC GCTCCGAGCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5062)	2	TREX2_+_152736011.23-P1P2	TREX2	P1P2	HAUS7_+_152736011.23-P1	HAUS7	P1	na	na	na	na	na	na
GGGCGGCGACGACTACT CAG (SEQ ID NO: 1917)	TATGAACCACTAAGGCGTccaccttgttgGGGCGGCGACGAC TACTCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3490)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGAGTA GTCGTCGCCGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5063)	2	TREX2_+_152735963.23-P1P2	TREX2	P1P2	HAUS7_+_152735963.23-P1	HAUS7	P1	na	na	na	na	na	na
GGCGCAGCCGGAACCGG CGA (SEQ ID NO: 1918)	TATGAACCACTAAGGCGTccaccttgttgGGCGCAGCCGGAA CCGGCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3491)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCCGG TTCCGGCTGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5064)	2	TREX2_-_152735804.23-P1P2	TREX2	P1P2	na	na	na	na	na	na	na	na	na
GCGGCCGAGCGGTTGCG ACG (SEQ ID NO: 1919)	TATGAACCACTAAGGCGTccaccttgttgGCGGCCGAGCGGT TGCGACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3492)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTCGCA ACCGCTCGGCCGCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 5065)	2	TRRAP_-_98476196.23-P1P2	TRRAP	P1P2	na	na	na	na	na	na	na	na	na
GGCAGGCCGCACCCACT TGG (SEQ ID NO: 1920)	TATGAACCACTAAGGCGTccaccttgttgGGCAGGCCGCACC CACTTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3493)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCAAGTG GGTGCGGCCTGCCCAACAAGGTGGACGCCTTAGTGGTTCAT A (SEQ ID NO: 5066)	2	TRRAP_-_98476108.23-P1P2	TRRAP	P1P2	na	na	na	na	na	na	na	na	na
GTGCGACGAGGGCTCGG CTG (SEQ ID NO: 1921)	TATGAACCACTAAGGCGTccaccttgttgGTGCGACGAGGGC TCGGCTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3494)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCAGCCGA GCCCTCGTCGCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5067)	2	TRRAP_-_98476208.23-P1P2	TRRAP	P1P2	na	na	na	na	na	na	na	na	na

GGCGGGACAACCACGAGCCG (SEQ ID NO: 1922)	TATGAACCACTAAGGCGTccaccttgttgGGCGGGACAACCACGAGCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3495)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGCTCGTGGTTGTCCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5068)	1	UBA2_+_34919314.23-P1P2	UBA2	P1P2	na	na	na	na	na	na	na	na	na
GGGACAACCACGAGCCGCGG (SEQ ID NO: 1923)	TATGAACCACTAAGGCGTccaccttgttgGGGACAACCACGAGCCGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3496)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCGGCTCGTGGTTGTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5069)	1	UBA2_+_34919311.23-P1P2	UBA2	P1P2	na	na	na	na	na	na	na	na	na
GTCCCGCCATGGCACTGTCG (SEQ ID NO: 1924)	TATGAACCACTAAGGCGTccaccttgttgGTCCCGCCATGGCACTGTCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3497)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGACAGTGCCATGGCGGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5070)	1	UBA2_-_34919351.23-P1P2	UBA2	P1P2	na	na	na	na	na	na	na	na	na
GCCGACCCTCGACTTCGGAG (SEQ ID NO: 1925)	TATGAACCACTAAGGCGTccaccttgttgGCCGACCCTCGACTTCGGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3498)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCGAAGTCGAGGGTCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5071)	1	UBE2A_-_11870854923-P1P2	UBE2A	P1P2	na	na	na	na	na	na	na	na	na
GGAGGCAGCGCGGTTCCTCT (SEQ ID NO: 1926)	TATGAACCACTAAGGCGTccaccttgttgGGAGGCAGCGCGGTTCCTCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3499)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGAGGAACCGCGCTGCCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5072)	1	UBE2A_-_118708567.23-P1P2	UBE2A	P1P2	na	na	na	na	na	na	na	na	na
GGGGAGGCGGAAGCACCCAG (SEQ ID NO: 1927)	TATGAACCACTAAGGCGTccaccttgttgGGGGAGGCGGAAGCACCCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3500)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGGGTGCTTCCGCCTCCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5073)	1	UBE2A_+_118708560.23-P1P2	UBE2A	P1P2	na	na	na	na	na	na	na	na	na
GGAGCATGTCGACCCCGGCC (SEQ ID NO: 1928)	TATGAACCACTAAGGCGTccaccttgttgGGAGCATGTCGACCCCGGCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3501)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGCCGGGGTCGACATGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5074)	1	UBE2B_-_13370730323-P1P2	UBE2B	P1P2	na	na	na	na	na	na	na	na	na
GGCCCGGAGGAGGCTCATGC (SEQ ID NO: 1929)	TATGAACCACTAAGGCGTccaccttgttgGGCCCGGAGGAGGCTCATGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3502)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCATGAGCCTCCTCCGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5075)	1	UBE2B_-_133707319.23-P1P2	UBE2B	P1P2	na	na	na	na	na	na	na	na	na
GGGAACCGCGCCCCGCTGAA (SEQ ID NO: 1930)	TATGAACCACTAAGGCGTccaccttgttgGGGAACCGCGCCCCGCTGAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3503)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTCAGCGGGGCGCGGTTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5076)	1	UBE2B_-_133707036.23-P1P2	UBE2B	P1P2	na	na	na	na	na	na	na	na	na
GAGGGAAGTCCCGAGACAAA (SEQ ID NO: 1931)	TATGAACCACTAAGGCGTccaccttgttgGAGGGAAGTCCCGAGACAAAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3504)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTTTGTCTCGGGACTTCCCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5077)	1	UBE2I_-_1359701.23-P1P2	UBE2I	P1P2	na	na	na	na	na	na	na	na	na
GGAGCGTAGCGGACAGGTGG (SEQ ID NO: 1932)	TATGAACCACTAAGGCGTccaccttgttgGGAGCGTAGCGGACAGGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3505)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACCTGTCCGCTACGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5078)	1	UBE2I_+_1359730.23-P1P2	UBE2I	P1P2	na	na	na	na	na	na	na	na	na
GGGACAGGTGGAGGACCGAG (SEQ ID NO: 1933)	TATGAACCACTAAGGCGTccaccttgttgGGGACAGGTGGAGGACCGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3506)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCGGTCCTCCACCTGTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5079)	1	UBE2I_+_1359721.23-P1P2	UBE2I	P1P2	na	na	na	na	na	na	na	na	na
GACGCACGAGTGGAAGTCCC (SEQ ID NO: 1934)	TATGAACCACTAAGGCGTccaccttgttgGACGCACGAGTGGAAGTCCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3507)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGGACTTCCACTCGTGCGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5080)	1	UBE2N_-_93835734.23-P1P2	UBE2N	P1P2	na	na	na	na	na	na	na	na	na
GCCTGCCCGCGACGGTGGTA (SEQ ID NO: 1935)	TATGAACCACTAAGGCGTccaccttgttgGCCTGCCCGCGACGGTGGTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3508)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTACCACCGTCGCGGGCAGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5081)	1	UBE2N_-_93835998.23-P1P2	UBE2N	P1P2	na	na	na	na	na	na	na	na	na
GTGCGTGAGGCGAGAGGAGC (SEQ ID NO: 1936)	TATGAACCACTAAGGCGTccaccttgttgGTGCGTGAGGCGAGAGGAGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3509)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCTCCTCTCGCCTCACGCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5082)	1	UBE2N_+_93835696.23-P1P2	UBE2N	P1P2	na	na	na	na	na	na	na	na	na
GAGCAACGCGCGGACCAACG (SEQ ID NO: 1937)	TATGAACCACTAAGGCGTccaccttgttgGAGCAACGCGCGGACCAACGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3510)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGTTGGTCCGCGCGTTGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5083)	1	UBE2T_-_202311067.23-P1P2	UBE2T	P1P2	na	na	na	na	na	na	na	na	na
GCAACAGTCAGAGGTCGCGC (SEQ ID NO: 1938)	TATGAACCACTAAGGCGTccaccttgttgGCAACAGTCAGAGGTCGCGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3511)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCGCGACCTCTGACTGTTGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5084)	1	UBE2T_+_202311076.23-P1P2	UBE2T	P1P2	na	na	na	na	na	na	na	na	na

GTGAGACCCTCCGGTCAA CAT3 (SEQ ID NO: 1939)	TATGAAGCACTAAG3GCGTcaacctgtgGGAG3AGCCTCCGC	TCAACAG:gttaagagtaaagtgGTAAAGTCTGCT9GTTGATA (SEQ ID NO: 3512)	CGGCAG393TCTCCCAACAAG93TGGACCGCCTTACTGTGTTCATA (SEQ ID NO: 5085)	UBE1T_-_20231983 23-P1P2	UBE1T	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAATCAGGAGCCCGGG CCC (SEQ ID NO: 1940)	TATGAACCACTAAGGCGTcaacctgtgGAAATCACGCAAGC	TCAACAGCAAGAAGTCTTCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 3513)	TCAACAGCAAGAAGTCTTACGCAAGCGTTAGCTCTTAAACCAGCCGG GCTCCGTGTGATTTCCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5086)	UBE2V2_+_49821037 23-P1P2	UBE2V2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGCCCGGGAAACCCAC CTG (SEQ ID NO: 1941)	TATGAACCACTAAGGCGTcaacctgtgGCGGCCCGCGAAC	TCAACAGCAAGAAGTCTTCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 3514)	TCAACAGCAAGAAGTCTTACGCAAGCGTTAGCTCTTAAACCAGCCTCG GTTCCGCGGTGCGCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5087)	UBE2V2_+_49821023-P1P2	UBE2V2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAATGGCCGGTCTCCACAA GGT (SEQ ID NO: 1942)	TATGAACCACTAAGGCGTcaacctgtgCGATCGGCGGTCTG	CACAAGGTgttaagagtaaagtgGTAAAGTCTGCTGGTTCATA (SEQ ID NO: 3515)	TCAACAGCAAGAAGTCTTACGCAAGCGTTAGCTCTTAAACCAGCTCG GAGACCCGCCCATCCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5088)	UBE2V2_+_49821034 23-P1P2	UBE2V2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAACGGCCCGGGTCTG AGG (SEQ ID NO: 1943)	TATGAACCACTAAGGCGTcaacctgtgGAAGCGCCGCGG	TCTTGAGgttaagagtaaagtgGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3516)	TCAACAGCAAGAAGTCTTACGCAAGCGTTAGCTCTTAAACCCTCAGAG CGCGTCCGCCGCTTCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5089)	UBN1_-_4997536 23-P1P2	UBN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCACGAGCCGAAGGG AGG (SEQ ID NO: 1944)	TATGAACCACTAAGGCGTcaacctgtgGGTCACGAGCCGGA	ACGGAAGGgttaagagtaaagtgGTAAAGTCTGCTGGTTCATA (SEQ ID NO: 3517)	TCAACAGCAAGAAGTCTTACGCAAGCGTTAGCTCTTAAACCCTCCCTT CGGCTCGTGAGCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5090)	UBN1_+_4997489 23-P1P2	UBN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTTGTCACGAGCGCGAA CGG (SEQ ID NO: 1945)	TATGAACCACTAAGGCGTcaacctgtgCGTTGTCACGAGC	CCAAAGGAgttaagagtaaagtgCGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3518)	TCAACAGCAAGAAGTCTTACGCAAGCGTTAGCTCTTAAACCCTCCTTG GCTCGTGACAACCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5091)	UBN1_+_4997492 23-P1P2	UBN1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGGCGGGACAGCTGGTG CGG (SEQ ID NO: 1946)	TATGAACCACTAAGGCGTcaacctgtgGCCGGGCGCGGG	CGGTGACGgttaagagtaaagtgCGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3519)	TCACCAGCAAGAAGTTTACGCAAGCGTTAGCTCTTAAACCCGCACCC AGCTGCCGCGGGCGCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5092)	UCHL5_+_19030438 23-P1P2	UCHL5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAACGTGTGGTTGACGA GCT (SEQ ID NO: 1947)	TATGAACCACTAAGGCGTcaacctgtgGGACGTGTGGTTGA	ACGAAGCTgttaagagtaaagtgCGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3520)	TCAACAGCAAGAACTTTACGCAAGCGTTAGCTCTTAAACAGCTCGGT CAAACCACACGTCCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5093)	UCHL5_+_19030482 23-P1P2	UCHL5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTGTGCTTCGACGAGCCT CGG (SEQ ID NO: 1948)	TATGAAGCACTAAGGCGTcaacctgtgCGTTGTGGTTGAGCG	AGCTCGGgttaagagtaaagtgCGTAAAGTCTGCTGGTTGATA (SEQ ID NO: 3521)	TCAGCAGCAAGAACTTTACGCAAGCGTTAGCTCTTAAACCCGGAGCT CGTCAACCACACGCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5094)	UCHL5_+_19030479 23-P1P2	UCHL5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAAACCCCCACCCGAGAT CCG (SEQ ID NO: 1949)	TATGAACCACTAAGGCGTcaacctgtgGAAACCCCACCCG	AGAATCCCgttaagagtaaagtgCGTAAAGTCTCCGTGGTTGA (SEQ ID NO: 3522)	TCAACAGCAAGAAGTTTACGCAAGCGTTAGCTCTTTAAACCCGGATCT CGGGTTGGGGTTTTCCCAACAAGGTCGAACCGCCTTAGTGGTTCATA (SEQ ID NO: 5095)	UHRF1_+_49103401 23-P1P2	UHRF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGCCGCGCCGGGACTA CAG (SEQ ID NO: 1950)	TATGAACCACTAAGGCGTcaacctgtgGCCCGCGCCGGG	ACTACAGgttaagagtaaagtgCGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3523)	TCAACAGCAAGAACTTTACGCAAGCGTTAGCTCTTAAACCTGTAGTC CCCGGCGGCCGGCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5096)	UHRF1_+_49103126 23-P1P2	UHRF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTGCTGCGGGAGCGCAAT AAG (SEQ ID NO: 1951)	TATGAAGCACTAAGGCGTcaacctgtgCGTTGCTGCGGAGC	CAATAAGgttaagagtaaagtgCGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3524)	TCAGCAGCAAGAAGTTTACGCAAGCGTTAGCTCTTAAAGCTTATTGG CTCCGCGAGCAGCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5097)	UHRF1_-_49103083 23-P1P2	UHRF1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCACACGTTGGGGAGCG CGG (SEQ ID NO: 1952)	TATGAACCACTAAGGCGTcaacctgtgGCCACGTTGGGG	AGCGCGCGgttaagagtaaagtgCGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3525)	TCACCAGCAGAAGTTTACGCAAGCGTTAGCTCTTAAACGCGCGCGCT CGCCAACGTGTGGCGCAACAAGTGGAACGGCGTTAGTGGTTCATA (SEQ ID NO: 5098)	UMC1_-_17643379 27-P1P2	UMC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGTCTGCGAGAGAGCACA CCC (SEQ ID NO: 1953)	TATGAACCACTAAGGCGTcaacctgtgAGCGTCGAGAGCA	CACACCCGgttaagagtaaagtgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3526)	TCACCAGCAGAAGTCTTACGCAAGCGTTAGCTCTTAAACGCGGTCTCG TCTCTCTCGCGACCGCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5099)	UMC1_-_17643380 23-P1P2	UMC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGCTCTGGCGAGAGCGCCTA CCT (SEQ ID NO: 1954)	TATGAAGCACTAAGGCGTcaacctgtgCGCTCGGGAGAGG	CCTACCTgttaagagtaaagtgCGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3527)	CACCAGCAGAAGTCTTACGCAAGCGTTAGCTCTTAAACGAGCTAGG CCTGTCCCGAGACGCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5100)	UMC1_-_17643364 23-P1P2	UMC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCCGAGCCCGGCCCTTC CAG (SEQ ID NO: 1955)	TATGAACCACTAAGGCGTcaacctgtgGCCCGAGCCCGGC	CTTCCAGgttaagagtaaagtgCGTAAAGTCTGCTGGTGATA (SEQ ID NO: 3528)	CACCAGCAGAAGTCTTACGCAAGCGTTAGCTCTTAAACGTTCGAACG GCCGTCGCTCCGCGGCCAACAAGGTCGAAGCGCTTAGTGGTTCATA (SEQ ID NO: 5101)	UNG_-_10953573 23-P1P2	UNG	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GCGCCACCCGGGAACCGGCG (SEQ ID NO: 1956)	TATGAACCACTAAGGCGTccaccttgttgGCGCCACCCGGGAACCGGCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3529)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCGGTTCCCGGGTGGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5102)	1	UNG_+_109535435.23-P1P2	UNG	P1P2	na	na	na	na	na	na	na	na	na
GGTCGCTTCCTGGCGGGGCT (SEQ ID NO: 1957)	TATGAACCACTAAGGCGTccaccttgttgGGTCGCTTCCTGGCGGGGCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3530)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGCCCCGCCAGGAAGCGACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5103)	1	UNG_+_109535520.23-P1P2	UNG	P1P2	na	na	na	na	na	na	na	na	na
GGAGCGGCTGGCGGCTTCGA (SEQ ID NO: 1958)	TATGAACCACTAAGGCGTccaccttgttgGgagcggctggcggcttcgagtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3531)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGAAGCCGCCAGCCGCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5104)	1	UPF1_-_18942807.23-P1P2	UPF1	P1P2	na	na	na	na	na	na	na	na	na
GGCCCGAGGGCCCTACCCGG (SEQ ID NO: 1959)	TATGAACCACTAAGGCGTccaccttgttgGGCCCGAGGGCCCTACCCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3532)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGGGTAGGGCCCTCGGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5105)	1	UPF1_-_18943013.23-P1P2	UPF1	P1P2	na	na	na	na	na	na	na	na	na
GTTCGAGGGGAGCTGAGGCG (SEQ ID NO: 1960)	TATGAACCACTAAGGCGTccaccttgttgGttcgaggggagctgaggcggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3533)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGCCTCAGCTCCCCTCGAACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5106)	1	UPF1_-_18942821.23-P1P2	UPF1	P1P2	na	na	na	na	na	na	na	na	na
GCCAGGAGCGCAACAAGCAG (SEQ ID NO: 1961)	TATGAACCACTAAGGCGTccaccttgttgGCCAGGAGCGCAACAAGCAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3534)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTGCTTGTTGCGCTCCTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5107)	1	USP1_+_62902791.23-P1P2	USP1	P1P2	na	na	na	na	na	na	na	na	na
GCCGAACGCCTGAACCGCTA (SEQ ID NO: 1962)	TATGAACCACTAAGGCGTccaccttgttgGCCGAACGCCTGAACCGCTAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3535)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTAGCGGTTCAGGCGTTCGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5108)	1	USP1_+_62902760.23-P1P2	USP1	P1P2	na	na	na	na	na	na	na	na	na
GCGCGCCAAGTTCCCCTCGG (SEQ ID NO: 1963)	TATGAACCACTAAGGCGTccaccttgttgGCGCGCCAAGTTCCCCTCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3536)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGAGGGGAACTTGGCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5109)	1	USP1_-_62902745.23-P1P2	USP1	P1P2	na	na	na	na	na	na	na	na	na
GACCCCCTGTCCCGCCCGGC (SEQ ID NO: 1964)	TATGAACCACTAAGGCGTccaccttgttgGACCCCCTGTCCCGCCCGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3537)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCGGGCGGGACAGGGGGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5110)	1	USP28_-_113746063.23-P1P2	USP28	P1P2	na	na	na	na	na	na	na	na	na
GGAGACCGGTGACCCGCGGC (SEQ ID NO: 1965)	TATGAACCACTAAGGCGTccaccttgttgGGAGACCGGTGACCCGCGGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3538)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCCGCGGGTCACCGGTCTCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5111)	1	USP28_+_113746234.23-P1P2	USP28	P1P2	na	na	na	na	na	na	na	na	na
GGCGGGAGACCGGTGACCCG (SEQ ID NO: 1966)	TATGAACCACTAAGGCGTccaccttgttgGGCGGGAGACCGGTGACCCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3539)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGGTCACCGGTCTCCCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5112)	1	USP28_+_113746238.23-P1P2	USP28	P1P2	na	na	na	na	na	na	na	na	na
GAATGTCACATCCGCATGGG (SEQ ID NO: 1967)	TATGAACCACTAAGGCGTccaccttgttgGAATGTCACATCCGCATGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3540)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCATGCGGATGTGACATTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5113)	1	USP7_-_9057778.23-P1P2	USP7	P1P2	na	na	na	na	na	na	na	na	na
GAGCGAGGAGAGCAAGCGGG (SEQ ID NO: 1968)	TATGAACCACTAAGGCGTccaccttgttgGagcgaggagagcaagcggggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3541)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCGCTTGCTCTCCTCGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5114)	1	USP7_-_9057818.23-P1P2	USP7	P1P2	na	na	na	na	na	na	na	na	na
GCTCCTCGGCGTCGTCGTCG (SEQ ID NO: 1969)	TATGAACCACTAAGGCGTccaccttgttgGCTCCTCGGCGTCGTCGTCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3542)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGACGACGACGCCGAGGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5115)	1	USP7_-_9057691.23-P1P2	USP7	P1P2	na	na	na	na	na	na	na	na	na
GAGAGCCCCGGGGACCGAAC (SEQ ID NO: 1970)	TATGAACCACTAAGGCGTccaccttgttgGAGAGCCCCGGGGACCGAACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3543)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTTCGGTCCCCGGGGCTCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5116)	1	UVSSA_+_1341057.23-P1P2	UVSSA	P1P2	na	na	na	na	na	na	na	na	na
GGCCGCGTCAGCGGTAGGTG (SEQ ID NO: 1971)	TATGAACCACTAAGGCGTccaccttgttgGGCCGCGTCAGCGGTAGGTGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3544)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCACCTACCGCTGACGCGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5117)	1	UVSSA_-_1341026.23-P1P2	UVSSA	P1P2	na	na	na	na	na	na	na	na	na
GTGGTTACGCTGCCGGGCGG (SEQ ID NO: 1972)	TATGAACCACTAAGGCGTccaccttgttgGTGGTTACGCTGCCGGGCGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3545)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCGCCCGGCAGCGTAACCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5118)	1	UVSSA_-_1341051.23-P1P2	UVSSA	P1P2	na	na	na	na	na	na	na	na	na

GCCAAAACCGAGCTGGTGAG (SEQ ID NO: 1973)	TATGAACCACTAAGGCGTccaccttgttgGCCAAAACCGAGCTGGTGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3546)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCACCAGCTCGGTTTTGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5119)	2		VPS72_+_151162629.23-P1P2	VPS72	P1P2	na	na	na	na	na	na	na	na	na
GCCCCGGACAAAGAAGAAGA (SEQ ID NO: 1974)	TATGAACCACTAAGGCGTccaccttgttgGCCCCGGACAAAGAAGAAGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3547)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCTTCTTCTTTGTCCGGGGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5120)	2		VPS72_+_151162397.23-P1P2	VPS72	P1P2	na	na	na	na	na	na	na	na	na
GGTGGCGGCGCAGTCTCGGT (SEQ ID NO: 1975)	TATGAACCACTAAGGCGTccaccttgttgGGTGGCGGCGCAGTCTCGGTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3548)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACACCGAGACTGCGCCGCCACCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5121)	2		VPS72_+_151162601.23-P1P2	VPS72	P1P2	na	na	na	na	na	na	na	na	na
GACTGTGGGAAGCGGAGTGA (SEQ ID NO: 1976)	TATGAACCACTAAGGCGTccaccttgttgGACTGTGGGAAGCGGAGTGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3549)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCACTCCGCTTCCCACAGTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5122)	1		WAPAL_+_88281542.23-P1P2	WAPAL	P1P2	na	na	na	na	na	na	na	na	na
GAGCGGAGTGACGGAGCGAG (SEQ ID NO: 1977)	TATGAACCACTAAGGCGTccaccttgttgGAGCGGAGTGACGGAGCGAGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3550)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCGCTCCGTCACTCCGCTCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5123)	1		WAPAL_+_88281533.23-P1P2	WAPAL	P1P2	na	na	na	na	na	na	na	na	na
GTGAGAGCCGAGAGAGGCGA (SEQ ID NO: 1978)	TATGAACCACTAAGGCGTccaccttgttgGTGAGAGCCGAGAGAGGCGAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3551)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTCGCCTCTCTCGGCTCTCACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5124)	1		WAPAL_-_88281250.23-P1P2	WAPAL	P1P2	na	na	na	na	na	na	na	na	na
GCGGGGCGGCCCGAGCACCG (SEQ ID NO: 1979)	TATGAACCACTAAGGCGTccaccttgttgGCGGGGCGGCCCGAGCACCGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3552)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCGGTGCTCGGGCCGCCCCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5125)	1		WEE1_+_9595254.23-P1P2	WEE1	P1P2	na	na	na	na	na	na	na	na	na
GCTCGGGCTGCGGACGGTGG (SEQ ID NO: 1980)	TATGAACCACTAAGGCGTccaccttgttgGCTCGGGCTGCGGACGGTGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3553)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCACCGTCCGCAGCCCGAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5126)	1		WEE1_+_9595305.23-P1P2	WEE1	P1P2	na	na	na	na	na	na	na	na	na
GCTTTCCGGCAGAGGCGGGG (SEQ ID NO: 1981)	TATGAACCACTAAGGCGTccaccttgttgGCTTTCCGGCAGAGGCGGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3554)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCGCCTCTGCCGGAAAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5127)	1		WEE1_+_9595268.23-P1P2	WEE1	P1P2	na	na	na	na	na	na	na	na	na
GCTGTGGGAGGCCTGAGGGG (SEQ ID NO: 1982)	TATGAACCACTAAGGCGTccaccttgttgGCTGTGGGAGGCCTGAGGGGgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3555)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCCCCTCAGGCCTCCCACAGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5128)	1		WHSC1_-_1873356.23-P1P2	WHSC1	P1P2	na	na	na	na	na	na	na	na	na
GGCGCGCGAGAGCCTCGGCC (SEQ ID NO: 1983)	TATGAACCACTAAGGCGTccaccttgttgGgcgcgcgAGAGCCTCGGCCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3556)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGGCCGAGGCTCTCGCGCGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5129)	1		WHSC1_-_1873146.23-P1P2	WHSC1	P1P2	na	na	na	na	na	na	na	na	na
GGGCGGGCGGCTCTCCGGAG (SEQ ID NO: 1984)	TATGAACCACTAAGGCGTccaccttgttgGGGCGGGCGGCTCTCCggaggtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3557)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACCTCCGGAGAGCCGCCCGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5130)	1		WHSC1_+_1873227.23-P1P2	WHSC1	P1P2	na	na	na	na	na	na	na	na	na
GGCCTGGGTAGCACGGTCCT (SEQ ID NO: 1985)	TATGAACCACTAAGGCGTccaccttgttgGGCCTGGGTAGCACGGTCCTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3558)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGGACCGTGCTACCCAGGCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5131)	1		WHSC1L1_-_38239671.23-P1P2	WHSC1L1	P1P2	na	na	na	na	na	na	na	na	na
GGGCCGAGGGGGCTGTGCAC (SEQ ID NO: 1986)	TATGAACCACTAAGGCGTccaccttgttgGGGCCGAGGGGGCTGTGCACgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3559)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGTGCACAGCCCCCTCGGCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5132)	1		WHSC1L1_+_38239708.23-P1P2	WHSC1L1	P1P2	na	na	na	na	na	na	na	na	na
GTCCTCAGTGCCGGCCGCCA (SEQ ID NO: 1987)	TATGAACCACTAAGGCGTccaccttgttgGTCCTCAGTGCCGGCCGCCAgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3560)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACTGGCGGCCGGCACTGAGGACCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5133)	1		WHSC1L1_+_38239493.23-P1P2	WHSC1L1	P1P2	na	na	na	na	na	na	na	na	na
GCGCGCTGCCGAGTACAAGC (SEQ ID NO: 1988)	TATGAACCACTAAGGCGTccaccttgttgGCGCGCTGCCGAGTACAAGCgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3561)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACGCTTGTACTCGGCAGCGCGCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5134)	1		WRN_+_30891350.23-P1P2	WRN	P1P2	na	na	na	na	na	na	na	na	na
GGGAGGCGCCGGCTTGTACT (SEQ ID NO: 1989)	TATGAACCACTAAGGCGTccaccttgttgGGGAGGCGCCGGCTTGTACTgtttaagagctaagctgCGTAAAGTCTGCTGGTGA (SEQ ID NO: 3562)	TCACCAGCAGACTTTACGCAGCTTAGCTCTTAAACAGTACAAGCCGGCGCCTCCCCAACAAGGTGGACGCCTTAGTGGTTCATA (SEQ ID NO: 5135)	1		WRN_-_30891364.23-P1P2	WRN	P1P2	na	na	na	na	na	na	na	na	na

GTTAGACTATAGATACC GG (SEQ ID NO: 1950)	TATGAAGCACTAAGATACCACCGgttaagagtcaagcgtCGTAAAGTTCGCTGCTGA (SEQ ID NO: 3553)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACCTCTGGGTA 1GTATAGTTGCTACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5136)	WRN_+_3681 97527.23-P1P2	WRN	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACATGACGATCCGGT CTC (SEQ ID NO: 1951)	TATGAACCACTAAGACGCTcaacctgtgGACATGACGATC (SEQ ID NO: 3554)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACGAGACGG GATGGTCATGTTCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5137)	XAB2_-_7994228.23-P1P2	XAB2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGACTCACGAACACAAG GTC (SEQ ID NO: 1952)	TATGAACCACTAAGACGCTcaacctgtgGACTCACGAACA (SEQ ID NO: 3555)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACGACTTTGT CAAGGTTCgttaagagtcaagcgtCGTAAAGTTCTGCTGGTTCATA (SEQ ID NO: 5138)	XAB2_-_7994376.23-P1P2	XAB2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCGCGGTCTCTCGCCCA TTA (SEQ ID NO: 1953)	TATGAACCACTAAGACGCTcaacctgtgGTCTCGCGTCTCTGTGG CCACATTAgttaagagtcaagcgtCGTAAAGTTCGCTGGTTGA (SEQ ID NO: 3556)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACTTAATGGG CCACAGCAGCCGGACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5139)	XAB2_-_7994239.23-P1P2	XAB2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTGGGCCAGAGATGG CGG (SEQ ID NO: 1954)	TATGAACCACTAAGACGCTcaacctgtgGAGTGGGCCAGAG ATGCGCGGgttaagagtcaagcgtCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3557)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACCCGCCAT CTCTTGTGCCCACTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5140)	XPA_+_1004 69564.23-P1P2	XPA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCTCGGAGTGGGCCAG AGA (SEQ ID NO: 1955)	TATGAACCACTAAGACGCTcaacctgtgGCCTCGGAGTGGG CCAGAGAgttaagagtcaagcgtCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3558)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACTCTCTGGG CCACTCCGAGGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5141)	XPA_+_1004 69570.23-P1P2	XPA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGAAGCTAGGTCCTCGG AGT (SEQ ID NO: 1956)	TATGAAGCACTAAGATCCTCGGgttaagagtcaagcgtCGTAAAGTTCGCTGGTGA (SEQ ID NO: 3559)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACACTCCGGA GGAACCTAGCTCCCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5142)	XPA_+_1004 64580.23-P1P2	XPA	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCAAGGGCGGGAGTCCAG ACG (SEQ ID NO: 1957)	TATGAACCACTAAGACGCTcaacctgtgGCAAGGGCGGGAGT ACTCCCTCGCCTGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3573)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACGCTCTTG 9880.23-P1	XPC_+_1421	XPC	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCAGACAACAAGCAACATG GCT (SEQ ID NO: 1958)	TATGAACCACTAAGACGCTcaacctgtgGCCAGACAACATG TGCTTGTCTGTGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 3574)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACAGCATGTT 6059.23-P1	XPC_+_1422	XPC	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCGCAGTTCGCGTCGC CCG (SEQ ID NO: 1959)	TATGAAGCACTAAGATCCTCGCGgttaagagtcaagcgtCGTAAAGTTCGCTGGTGA (SEQ ID NO: 3575)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACCCGCGAC GACGAACTGCGCACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5145)	XPC_-_14220037.2-3-P1	XPC	P1	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCTGTGTAAAGACCACAG AGA (SEQ ID NO: 2000)	TATGAACCACTAAGACGCTcaacctgtgGCTGTGTAAAGACC AGAGCGAgttaagagtcaagcgtCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3573)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACGTCGTGTG GTCTTTACACAGCGCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5146)	XPC_-_14212071.2-3-P2	XPC	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTCCCTGTGTGTAAAGAC GAC (SEQ ID NO: 2001)	TATGAACCACTAAGACGCTcaacctgtgGGTCCCTGTGTGTAAA GACGCACgttaagagtcaagcgtCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3574)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACGTGGTGTGT TAAACAGGGCACCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5147)	XPC_-_14212067.2-3-P2	XPC	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTACGCTTGAATGCGAACA CT (SEQ ID NO: 2002)	TATGAACCACTAAGACGCTcaacctgtgCGTTACGCTTGAATGC AACACTgttaagagtcaagcgtCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3575)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACGATGTTCC ATTCAAGCGTAACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5148)	XPC_+_1421 20194.23-P2	XPC	P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCTGCTCGAGTGAGACG CA (SEQ ID NO: 2003)	TATGAACCACTAAGACGCTcaacctgtgGGCTGCTCGAGTGA GACGCAGgttaagagtcaagcgtCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3576)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACGTGTGTGT CACGTGGAGCAGTCCGAAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5149)	XRCC1_-_44079688.2-3-P1P2	XRCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGTTAGAGTATGTGGTGTCC GAG (SEQ ID NO: 2004)	TATGAACCACTAAGACGCTcaacctgtgGGTTAGAGTATGTGG TCCGGAGgttaagagtcaagcgtCGTAAAGTTCTGCTGGTGA (SEQ ID NO: 3577)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACCTCGGAC TCCATACTCTACCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5150)	XRCC1_-_44079688.2-3-P1P2	XRCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTTCTTCGAGACACCAAT CC (SEQ ID NO: 2005)	TATGAACCACTAAGACGCTcaacctgtgGTTCTTCGAGACAC CAATCCgttaagagtcaagcgtCGTAAAGTTCGCTGGTGA (SEQ ID NO: 3578)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACGTAGTGGG TGTCTGTGAAGAACCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5151)	XRCC1_+_4079843.23-P1P2	XRCC1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GAGTCTTCGGCGACTTTGGT GAA (SEQ ID NO: 2006)	TATGAACCACTAAGACGCTcaacctgtgGAGTCTTCGGCGACT TTGGTGAAgttaagagtcaagcgtCGTAAAGTTCGCTGCTGA (SEQ ID NO: 3579)	TCACCAGCAGAACTTTAGCGCAGCTTAGCTCTTAAACTTCACCGAA CTGCGCGACGAGTCCAACAAGGTCGACGCCTTAGTGGTTCATA (SEQ ID NO: 5152)	XRCC2_+_1523732.23-P1P2	XRCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGATGTGTAGTCGCTTCC ATT (SEQ ID NO: 2007)	TATGAACCACTAAGGCGTcaacctgtggGGATGTGTAGTGC	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACATGGAAG GCACTACACATCCCAACAAGTGTGGACCGCTTAGTGCTTCATA (SEQ ID NO: 5153)					XRCC2_+_1 82373143.23-P1P2	XRCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCACTACACATCGCCC CGA (SEQ ID NO: 2008)	TATGAACCACTAAGGCGTcaacctgtggGCACTACACATCG	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACTCGGGGC GATGTGTAGTGCCCAACAAGTGTGGACCGCTTAGTGCTTCATA (SEQ ID NO: 5154)					XRCC2_-_ 19237317 4.23-P1P2	XRCC2	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGAGAGGAGCACGA CGG (SEQ ID NO: 2009)	TATGAACCACTAAGGCGTcaacctgtggGCGTAAAGTCTGTGTGTGA	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCCGTCGT GCTCCTCTCGGGCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5155)					XRCC3_+_1 94181786.23-P1P2	XRCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGAACCCCGCGGGAAG GTG (SEQ ID NO: 2010)	TATGAACCACTAAGGCGTcaacctgtggCGCGGAAACCCGCGTG	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCCACCTCCC GCCGGTTCCCGCCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5156)					XRCC3_+_1 94181784.23-P1P2	XRCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCAGTGCGGAACCCGC GGG (SEQ ID NO: 2011)	TATGAACCACTAAGGCGTcaacctgtggGGAGTGCGGAACCCGC	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCCCGGGAG GTTCCGCACTCCCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5157)					XRCC3_+_1 94181788.23-P1P2	XRCC3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCGGGAAGTAGAGTCACG GAG (SEQ ID NO: 2012)	TATGAACCACTAAGGCGTcaacctgtggGCGGAAGTAGAGT	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCTCCGTG ACTCTACTTCCGCCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5158)					XRCC4_-_ 82373464.2 3-P1P2	XRCC4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAGGTAGGATCCCGAA GTG (SEQ ID NO: 2013)	TATGAACCACTAAGGCGTcaacctgtggGGAGGTAGGATCC	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACCACCTCCG GATCCTACCTCCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5159)					XRCC4_-_ 82373465.2 3-P1P2	XRCC4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGATACCGGAAAGTAGAG TGA (SEQ ID NO: 2014)	TATGAACCACTAAGGCGTcaacctgtggGATACCGGAAAGT	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACTGACTCTA CTTCCGGTATCCCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5160)					XRCC4_-_ 82373469.2 3-P1P2	XRCC4	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCACCATGTTTGCCGGTC CTC (SEQ ID NO: 2015)	TATGAACCACTAAGGCGTcaacctgtggGCACCATGTTTGC	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACGAGAACC GGCAACATGGTGTGCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5161)					XRCC5_+_2 16974143.23-P1P2	XRCC5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCGCTTTCGCTCGCTTCT CC (SEQ ID NO: 2016)	TATGAAGCACTAAGGCGTcaacctgtggCGCGCTTTGCGTCGC TTCTTCCgttaagagctaagcgtgCGTAAAGTCTGCTCGGTGA (SEQ ID NO: 5162)						XRCC5_+_2 16974121.23-P1P2	XRCC5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGAACCAAAGCGCCGTGAG GAC (SEQ ID NO: 2017)	TATGAACCACTAAGGCGTcaacctgtggGGAACCAAAGCGCC	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTTAAAGGTCCTCA GGCGCTTTGGTCCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5163)					XRCC5_-_ 21697415 4.23-P1P2	XRCC5	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GATAACGGGCCCGCTTAC CTT (SEQ ID NO: 2018)	TATGAACCACTAAGGCGTcaacctgtggGATAACGGGCCCGC	TCACCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACAAGGTAA GCCGGGCCGTTATCCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5164)					XRCC6_+_4 20173423-P1P2	XRCC6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GCCCGGTTACCTTTGGC GCA (SEQ ID NO: 2019)	TATGAAGCACTAAGGCGTcaacctgtggCGCCCGCTTACCTTT	TCAGCAGCAGAGACTTTAGCGCAGGTTAGCTCTTAAACTGCGCCA AAAGGTAAGCGGGGCCAACAAGTGTGGACCGCTTAGTGGTTCATA (SEQ ID NO: 5165)					XRCC6_+_4 20173388.23-P1P2	XRCC6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTCCCGTGCGGCCAAACGGT AAGG (SEQ ID NO: 2020)	TATGAACCACTAAGGCGTcaacctgtggGTCCCGTGCGGCCAAA AAGGTAACgttaagagctaagcgtgCGTAAAGTCTGCTCGGTGA (SEQ ID NO: 5166)						XRCC6_-_ 42017356.2 3-P1P2	XRCC6	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGCACCGGGCCCGGGGAG GGAA (SEQ ID NO: 2021)	TATGAACCACTAAGGCGTcaacctgtggGGCACCGGGCCCGG GAGGGAAgttaagagctaagcgtgCGTAAAGTCTGCTCGGTGA (SEQ ID NO: 5167)						YBX3_-_ 10975981.2 3-P1P2	YBX3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GTGCGCGCGACCGTCCC GGGA (SEQ ID NO: 2022)	TATGAAGCACTAAGGCGTcaacctgtggCGCGGTGCGACCCGG CCCGGGAgttaagagctaagcgtgCGTAAAGTCTGCTCGGTGA (SEQ ID NO: 5168)						YBX3_-_ 10975976.2 3-P1P2	YBX3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na
GGGCTGGGCCCGCCCCA CTA (SEQ ID NO: 2023)	TATGAACCACTAAGGCGTcaacctgtggTGGGCTGGGCCCGC CCCACTAgttaagagctaagcgtgCGTAAAGTCTGCTCGGTGA (SEQ ID NO: 5169)						YBX3_-_ 10975973.2 3-P1P2	YBX3	P1P2	na	na	na	na	na	na	na	na	na	na	na	na	na	na

GGAAGGAGGGCGAGTTA GTT (SEQ ID NO: 2024)	TATGAACCACTAAGGCGTcaacttgtgGAAGGAGGGCGA	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAAGAAGTAACT CGCCCTCCTTCCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5170)	1	YEAT54_+_97536174.23-P1P2	EAT54	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GGAGAAGCGCGCGAAAA GAG (SEQ ID NO: 2025)	TATGAACCACTAAGGCGTcaacttgtgGAGAAGCGCGCG	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACCTCTTTTC GCGCGCTTTGTTCCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5171)	1	YEAT54_+_97536179.23-P1P2	EAT54	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GGAGTTACTTGGGCTCA CGG (SEQ ID NO: 2026)	TATGAACCACTAAGCGCGTcaacttgtgGAGTTACTTGCG	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACCCGTGAG CTCACGGGTTAACTCCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5172)	1	YEAT54_+_97536834.23-P1P2	EAT54	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GAGGGCTGTCGCCGAGGG CCA (SEQ ID NO: 2027)	TATGAACCACTAAGGCGTcaacttgtgGAGGGTCGCCC	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACTGGCCTC GGGCCAGCAGCCTCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5173)	1	YY1_+_10077823.23-P1P2	YY1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GGCCGGGCCGAGAGCGAG GTG (SEQ ID NO: 2028)	TATGAACCACTAAGGCGTcaacttgtgGGCCGGGCCGAG	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACCACTCTGG CCGCGGCCCGCCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5174)	1	YY1_-_100706185.23-P1P2	YY1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GGCGGGCTGGAGAGCGAT CCATCAT (SEQ ID NO: 2029)	TATGAACCACTAAGGCGTcaacttgtgGGCGGCTGCAGAG	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACATGATCGG CTGTGCAGCCGGCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5175)	1	YY1_+_10057665824.23-P1P2	YY1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GAGAGTGCGGGCGAGAG ACG (SEQ ID NO: 2030)	TATGAACCACTAAGGCGTcaacttgtgGAGAGTGCGGGCG	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAAGCGTCTCG GCCCGCACTCTCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5176)	1	ZNHIT1_-_100880870.23-P1P2	ZNHIT1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GAGTTGTGTTGTGCCAAT 393 (SEQ ID NO: 2031)	TATGAACCACTAAGGCGTcaacttgtgGAGTTGTGTTGTGC	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACCCATTGG CACAACACAGACTCCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5177)	1	ZNHIT1_-_100881482.23-P1P2	ZNHIT1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GGAACAGTTGTGTTGTGC CAA (SEQ ID NO: 2032)	TATGAACCACTAAGGCGTcaacttgtgGGAACAGTTGTG	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACTTGGCACA ACACAACTGTCCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5178)	1	ZNHIT1_-_100881479.23-P1P2	ZNHIT1	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GACGTCGAGGCTGCAACT GT (SEQ ID NO: 2033)	TATGAAGCACTAAGGCGTcaacttgtgGACGTCGAGGCTGC	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACACGAGTT GGAGCCTCGAGGTCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5179)	1	ZFAN53_-_136286746.23-P1P2	ZFAN53	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GCACGGCGACAAGTGAA CGG (SEQ ID NO: 2034)	TATGAACCACTAAGGCGTcaacttgtgGCACGGCGACAAG	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACCGTTCAC TTGTCCGCGGTCGCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5180)	1	ZFAN53_-_136288444.23-P1P2	ZFAN53	P1P2	na	na	na	na	na	na	na	na	na	na	na	na
GGCGGGGAGTTGGAGATT CTG (SEQ ID NO: 2035)	TATGAACCACTAAGGCGTcaacttgtgGGCGGGAGTTGGA	TCACCAGCAGAGCTTAGCGAGCTTAGCTCTTAAACCAGAATCT CCAACTCCCGCCCAACAAGGTCGCAGCCTTAGTGGTTCATA (SEQ ID NO: 5181)	1	ZFAN53_+_136288107.23-P1P2	ZFAN53	P1P2	na	na	na	na	na	na	na	na	na	na	na	na

Table 6

	% sequencing reads with the intended +6 G to C edit in Repair-seq screens													
CRISRPi_sgRNA	K562 PE2 rep1	K562 PE2 rep2	K562 PE3+50 rep1	K562 PE3+50 rep2	HeLa PE2 rep1	HeLa PE2 rep2	HeLa PE3+50 rep1	HeLa PE3+50 rep2	HeLa PE3-50 rep1	HeLa PE3-50 rep2
ABRAXAS1_1	4.213310184	4.968476882	15.81769298	13.57270033	8.425562621	8.510656136	14.10582728	15.33239528	2.155470993	2.219974716
ABRAXAS1_2	4.200789601	4.769450073	16.33515095	13.68144765	8.287835769	8.826508808	13.90054067	15.94840646	2.023478436	2.185041705
ABRAXAS1_3	4.860138916	5.092866827	16.50551832	15.32124618	8.551238175	9.425665469	13.68919829	15.53391933	2.068269716	2.312933301
ACTL6A_1	4.975693452	6.671449067	20.30449827	19.03016164	8.850735057	8.450632305	13.89806788	15.80386701	2.038010664	2.020325327
ACTL6A_2	6.981803668	7.509400374	23.44918237	21.04264396	8.748845633	8.703612621	13.8876764	15.36776097	2.058519072	1.837442296
ACTL6A_3	6.87618857	7.960525114	24.71466198	22.84283738	8.819497437	9.204311109	13.77210565	17.24247097	2.666883044	2.265589091
ACTR5_1	4.942822435	5.895665925	15.50564454	6.939668646	7.125763267	11.12456282	11.4469701	1.301423607	1.48265947
ACTR5_2	5.186745859	5.797917253	18.06768601	15.32659105	6.162155407	6.801897963	11.07541871	12.03590778	1.566046746	1.312169312
ACTR5_3	4.272697571	5.681364842	17.58672284	14.63821647	7.596717987	7.79707657	12.16855773	13.6221388	1.877992015	1.693659569
ACTR8_1	5.976114803	6.431369427	20.23473884	18.06904797	7.658599805	8.044996691	12.66664309	13.90093882	1.872300397	1.403813393
ACTR8_2	6.352665505	6.76849325	20.54085207	18.50177069	7.166125619	7.006941611	11.7836441	12.8172679	1.806502889	1.67174226
ACTR8_3	6.176417122	6.487348756	20.15458557	17.6063	7.093032548	7.246166044	12.23555469	13.10499337	1.91007611	1.857338137
AICDA_1	4.444412284	5.05891692	15.62242617	13.83301668	9.00712641	9.209726435	13.72161982	15.71105717	2.062261228	2.156266823
AICDA_2	4.232903908	4.939801223	15.74766852	14.6457433	8.360723898	9.006266152	13.93234085	14.95615963	2.253644872	2.117518146
AICDA_3	4.229195088	5.149806852	16.65204066	14.18223437	8.361555154	8.59031312	13.69251328	15.64045393	2.201483426	2.179608213
ALKBH2_1	4.291606075	4.588767936	15.72361706	13.09856896	6.686637893	7.240132864	11.90354497	13.46530373	1.523449466	1.628202341
ALKBH2_2	3.864307917	4.600497963	14.39273685	12.91034403	6.780484965	6.75237065	11.19984441	13.29425602	1.687355711	1.7310054 72
ALKBH2_3	3.747244673	4.637762763	15.07660447	14.00259457	6.449735969	6.606814371	11.56047168	12.16788473	1.620830345	1.868883682
ALKBH3_1	4.081644424	4.68310249	15.26041665	13.5903842	8.733750602	8.881180407	14.34114206	15.62164941	2.192977454	2.03410006
ALKBH3_2	4.325463221	4.724403492	15.34089138	14.70515506	8.618054402	8.8253276	14.41883691	14.97357855	2.088886846	1.989764194
ALKBH3_3	3.905400278	4.704639816	15.06308688	14.6383443	8.460515441	8.48350341	14.07673885	15.86838932	2.531503008	2.014432405
ALYREF_1	4.612957624	4.965001731	15.64563721	13.74276676	8.203616587	8.422763845	13.59607528	15.14332756	2.273339483	2.129975921
ALYREF_2	4.816739222	4.907265672	16.03588694	14.26701788	7.239578788	7.484078433	13.0249334 6	14.09951053	2.031311337	1.848664623
ALYREF_3	4.868778281	5.373134328	16.08496643	14.71949545	7.126537096	7.513622647	12.5232007 1	14.37484791	1.767442652	2.020941642
APEX1_1	4.361202465	5.140916016	16.56543429	14.53601338	8.689856162	8.594046865	14.53996538	16.02200863	2.330065769	2.090021976
APEX1_2	4.357098931	5.28657554	15.75612364	14.37785352	8.659725951	8.617963635	14.37610323	15.82157137	1.865857872	1.782804089
APEX1_3	4.401026773	5.271044362	15.74321685	14.885242255	7.90419991	8.358279976	13.3896785 1	14.93339415	2.191465769	2.096292925
APEX2_1	4.013823063	4.902403427	16.02529949	13.17783272	8.442789076	8.865992999	14.1866424	15.31340075	2.07540611	1.966209718
APEX2_2	4.643287176	4.816139038	16.93696573	14.78079962	9.009792697	9.25354814	14.40300816	15.21730882	2.151751343	1.987540789
APEX2_3	4.433329012	5.155898313	16.35855135	13.59423111	8.320872866	8.762911105	14.3941622 1	15.23240669	1.969789442	2.034296225
APLF_1	4.131870314	4.507293987	15.80009769	14.32931456	8.079451755	8.434003786	13.83778745	15.73508047	2.398023594	2.162063953
APLF_2	4.264840563	5.018646761	15.90978661	15.2609492	8.641028933	8.786307981	14.2251319 5	16.35307859	2.456476987	2.148504108
APLF_3	4.457435581	4.960435698	16.55230866	14.70051578	8.635312574	8.279020087	14.05517063	15.53472915	2.189134811	2.038399001
APOBEC1_1	4.272839968	5.008864158	15.71959905	13.83452162	8.274293411	8.779351172	13.94246817	15.00419922	2.261754892	2.127911101
APOBEC1_2	4.324437907	4.522572699	15.89791961	14.11416269	8.460317971	8.774979605	13.56102798	15.70135065	1.879415329	1.893290902
APOBEC1_3	4.127040038	4.58531632	17.69771885	15.35839517	8.704169588	8.426441027	13.63822028	15.09055798	2.173115716	1.819538944
APOBEC2_1	4.090308405	4.581092653	15.40528545	14.29789838	8.014244825	8.234908675	14.45733417	15.39886343	2.326167654	1.990986365
APOBEC2_2	3.785581272	4.819406236	14.78805464	14.22209824	8.864319748	8.670270761	14.99008794	15.35688233	2.603616891	2.113344942
APOBEC2_3	4.461194491	5.067140633	15.41780997	13.66078913	8.970849488	9.221794048	14.22361606	15.22454572	2.093412052	1.984456535
APOBEC3A_1	4.111172103	4.682566582	15.46126277	14.47460179	8.669331738	8.404532855	13.66111962	14.73805326	2.362195452	2.210654661
APOBEC3A_2	4.532344458	4.932850742	15.70525032	14.30145034	8.503513976	8.767127725	13.82225114	15.18586936	2.140180926	2.387404568

APOBEC3A_3	4.498522834	5.215812495	15.61737119	14.77939578	8.316210424	8.517699115	14.0562095	15.62857984	1.732929494	2.068788384
APOBEC3B_1	3.76861181	4.364728502	14.17351603	12.72210568	7.420149859	7.238675184	12.5235051	13.68100327	2.349125859	2.446385139
APOBEC3B_2	4.350228632	5.279426834	15.46914053	14.78061357	9.326365139	9.898065691	14.0856257	16.21589872	2.340145873	2.456535809
APOBEC3B_3	4.11043247	4.776889428	15.87207636	13.53758338	8.749515392	8.722032242	14.47560292	15.62782239	2.171852174	2.141307424
APOBEC3C_1_APOBEC3D_1	4.305863347	4.812648816	15.63569618	14.30766384	8.382369816	8.60706065	14.3945943	15.76111819	2.016226893	2.263595207
APOBEC3C_2_APOBEC3D_2	4.307669573	5.070495877	16.12388111	14.0370062	8.500378779	8.373117094	13.9021290	15.17251328	2.341469092	1.979218938
APOBEC3C_3_APOBEC3D_3	4.322436836	5.059107029	15.5722661	13.79630349	8.007379623	8.569247639	13.941005	15.7840437	2.17053743	2.161064529
APOBEC3F_1	4.218834194	4.925157245	16.19562931	14.6031746	8.307469781	8.872993023	14.57280646	15.69566576	2.053915276	2.337166086
APOBEC3F_2	4.208029375	4.974702453	15.20684832	12.86690093	8.333424246	9.037837562	13.71998241	15.23345967	2.159803426	2.149056983
APOBEC3F_3	4.354405998	4.59059429	16.37806366	14.75520854	8.301500091	8.293542204	14.47317069	15.73472306	2.065015799	2.15493179
APOBEC3G_1	4.06957981	4.841192485	15.58189705	13.98504569	8.030488243	8.203733184	13.7564092	14.98904373	2.405102451	2.105164476
APOBEC3G_2	4.493942946	5.054056156	15.60362336	13.76205247	8.186157302	8.003692691	13.8450169	14.9281406	2.110502261	2.071064729
APOBEC3G_3	4.261887905	4.849398901	16.47500492	13.79705336	8.798015513	8.781763627	14.1682156	15.34600117	2.156997866	2.405666265
APOBEC3H_1	4.240547548	4.866767535	16.62287959	14.05132261	8.716394322	8.924176495	13.9386778	15.8162238	2.465036033	2.23802502
APOBEC3H_2	4.621110182	4.790176166	15.95819535	13.47619604	8.544135633	8.8099631	13.8575729	15.59665344	2.166634597	2.12697166
APOBEC3H_3	4.463210541	4.710321761	14.90779563	14.0557006	8.999074838	8.977815619	13.9962321	15.49758579	2.00311019	1.92714688
APOBEC4_1	4.374224718	4.90567968	15.11271265	13.90219167	9.053572382	8.970313607	14.5094887	16.20378741	2.089635149	2.138207462
APOBEC4_2	4.694787631	5.223624475	16.97254138	15.51704766	8.641367369	9.063587668	13.8562895	15.52798657	2.272124062	2.006389395
APOBEC4_3	4.595845824	4.951335364	16.28696784	14.87331741	8.596217208	9.062298076	13.8631740	15.57789372	2.325098394	2.174569162
APTX_1_1	4.032743279	5.012048193	15.82716658	13.56517519	8.129970199	8.802727962	14.1336947	15.55383166	2.109111361	1.872885987
APTX_1_2	4.526061536	5.139123042	16.30093858	14.8731578	8.338769585	8.833115453	13.7517314	15.20526263	2.097367676	1.913004835
APTX_1_3	4.357231115	4.763689841	16.07418019	14.3515158	8.161156488	8.437413267	13.974427	15.41327286	2.217385264	1.967446385
APTX_2_1	3.933106636	4.860250043	15.92059362	14.2774636	8.903041034	9.28481157	14.46713213	16.01827179	2.519449624	1.973023549
APTX_2_2	4.522706618	4.983766777	17.26770747	13.96430012	8.915189834	9.674803664	14.29816407	16.36683015	2.163573946	2.47746926
APTX_2_3	4.380492781	5.226671109	16.24334571	14.77568154	9.353641867	9.358580351	14.67603675	16.16803842	2.166751381	2.108059553
AQR_1	5.25	5.971433488	16.39893384	14.91200503	7.334376778	8.1724031866	13.7138806	15.34922718	2.254334133	2.281523519
AQR_2	3.935185185	4.676836612	14.08593092	13.78910776	7.221844361	8.130906139	12.6330298	14.74816773	2.122349873	2.046283206
AQR_3	3.835908364	4.225601132	12.26736973	9.764385951	7.559309828	8.101673212	13.94092664	14.3376736	2.354121384	2.266677551
ARID1A_1	3.944027087	4.475568495	14.89590363	14.01205566	8.189082956	8.031302356	14.0591693	15.27478518	2.106337862	1.551266586
ARID1A_2	4.342280329	4.58088221	15.17663852	13.46282952	8.082114736	8.467298968	13.7478916	15.02035615	2.233706069	2.040689967
ARID1A_3	3.710143755	4.48779229	14.91246072	13.46433241	7.996886629	7.879454578	13.4511615	14.93135281	2.056031432	1.994527509
ARID1B_1_1	4.30735561	5.192629816	16.75760729	14.40659186	7.836085393	7.812126317	13.5330636	15.06651925	2.006381855	2.037944041
ARID1B_1_2	4.261901865	4.918236005	16.36486673	14.41441411	8.768555159	9.129994106	14.5888372	15.63349647	2.455515211	2.337136593
ARID1B_1_3	4.383982356	5.092166128	15.18319007	13.97976613	8.727091039	9.219119089	14.4533063	15.03162197	1.995061466	2.379222328
ARID1B_2_1	4.010473328	4.786840653	15.46574686	13.78798423	8.140355699	8.529239768	13.5807171	15.25468318	2.088022797	2.226081119
ARID1B_2_2	4.119793953	4.722620756	15.01092029	13.3432975	8.480105922	8.19529863	13.8617341	15.21044108	2.147818779	2.142998567
ARID1B_2_3	4.156613292	5.34152417	17.37919921	14.98609524	8.270553813	9.134662318	14.5795145	15.38241623	2.196884765	1.764245572
ARID2_1	4.238565489	4.847771003	15.96118354	12.99578756	8.664751531	9.132664163	14.6751853	15.97206803	2.139784946	2.137903566
ARID2_2	4.933237397	5.284084265	16.93270009	14.03095608	11.35446218	10.92938616	16.5997322	18.18190658	2.765325256	2.883035967
ARID2_3	4.097459181	5.291043173	15.53746161	15.26869121	11.19085442	11.38565211	17.5933214	18.49096163	3.076097857	2.874213462
ASCC3_1	7.194185188	6.225581574	20.02804189	16.59962523	6.396299571	7.499152203	12.1789953	13.35685464	2.050496904	2.080735661
ASCC3_2	6.105966672	6.305468684	19.26426702	17.48386355	6.966825666	7.488877904	12.550234	13.54886036	2.13033851	1.865892916
ASCC3_3	5.925372088	6.225637267	19.30772148	18.91926674	7.103200896	7.872008464	12.4744412	13.71826109	1.996550046	2.10153234
ASF1A_1	4.643017039	5.300409195	17.74608276	15.37988291	7.715739989	8.111434223	13.3562555	14.40655195	2.327686707	2.13207294
ASF1A_2	4.834403234	5.294583071	17.19697966	15.0631411	7.923502949	8.11572951	13.3758790	14.37876376	2.265665628	2.072394626
ASF1A_3	4.839787976	5.593180239	17.96347617	15.51177736	8.029232916	7.899281053	13.5409765	14.27875316	2.036561601	1.995638316

ATM_1	4.540134749	5.2257411093	18.150274741	14.849716144	8.4152901953	8.6145320075	14.161206961	14.951880971	3.117240951	2.539071835			
ATM_2	4.515911857	4.992240895	17.762835791	15.047804091	8.4198446871	8.0509112821	13.980937361	15.45164471	2.439865741	2.470302596			
ATM_3	4.086291539	4.663480257	17.7401953	15.268771961	7.9767545161	8.1710728141	13.793593291	15.649704121	3.27451879	3.151573263			
ATP13A2_1	3.991159506	4.41628107	16.073624541	13.265771581	7.6266031331	7.9637431981	13.894787311	15.378459681	2.441132139	1.862459522			
ATP13A2_2	4.328758145	5.026994674	15.63120681	13.723421431	8.5155363361	8.0389539541	14.315066641	15.267105671	2.201292879	2.069545504			
ATP13A2_3	4.677242668	5.306670298	16.207427691	14.029239661	8.4658093291	8.8256565571	13.955054141	15.43881671	2.031314664	2.305350152			
ATP13A5_1	4.344795965	4.949662151	15.41405061	13.446295351	8.1650410411	8.5311487391	13.088334541	15.397812571	2.205346691	2.525735095			
ATP13A5_2	4.262394196	5.156301812	16.514933321	14.517228651	8.4701122221	8.4311632871	14.357120671	14.632459421	2.434099761	2.455594941			
ATP13A5_3	6.029761575	5.963285024	17.623867561	17.003587111	8.7899586361	9.1733221971	14.590969461	16.223392491	2.328879491	2.002649245			
ATR_1	6.202739147	6.36737557	15.749029341	14.751149311	9.5424888011	9.6015968711	14.45153451	16.244397791	2.174600079	2.143193941			
ATR_2	5.749285074	6.816118628	15.561273721	13.858415581	8.6938977721	9.1192386011	14.27142281	15.97316511	1.955492331	2.094404971			
ATR_3	5.877058294	6.413639131	15.751726491	13.420026011	9.5470228131	9.7480802171	14.515414151	15.508326381	2.293502591	2.207797411			
ATRIP_1	4.72609005	5.736901621	16.982035141	15.103168571	9.1576539561	9.9590244071	14.595842121	16.237706471	2.279569891	1.946632541			
ATRIP_2	4.772571713	5.789286491	16.454500811	15.591658761	8.7973402091	8.8020689931	13.979545111	15.643850221	2.107586231	1.982133821			
ATRIP_3	5.301309621	6.021232531	17.449731741	16.845941981	9.3775199731	9.2551428691	14.310626371	15.996181421	2.076490391	2.145678511			
ATRX_1	5.199345851	5.339556741	16.307222391	17.086207411	9.5924980641	9.4979435541	15.727687281	17.433644731	2.161307761	2.363097861			
ATRX_2	4.860741251	5.329475111	17.490991111	15.754215621	8.2360566411	8.8054541191	14.399976631	15.925860311	2.448362151	2.625820561			
ATRX_3	4.969696971	4.767981471	18.029847221	16.549007511	8.3575748951	9.1893902131	14.536918111	15.655794701	2.515079511	2.174997061			
BABAM1_1	4.267543651	5.215555241	15.849862721	14.715	8.5029483661	8.6681819131	14.675100341	15.175896271	2.145922741	2.425509261			
BABAM1_2	4.472600201	5.153004811	15.842063791	14.317107371	8.3866414671	8.8335752541	13.865546681	15.224806951	2.356932561	2.228857181			
BABAM1_3	4.350087571	4.901665951	15.453182861	13.768063061	8.1728202161	8.3057144381	14.216206561	15.779549441	1.947878151	2.18840601			
BABAM2_1	4.110170211	4.803647721	15.971099421	14.746741681	8.5793820771	9.1653687291	14.214738041	16.237191681	2.445652171	2.387743331			
BABAM2_2	4.764286901	5.285751189	15.911139731	14.624252631	9.0187670551	9.1694358431	14.982414631	15.786742841	2.478047491	2.163638561			
BABAM2_3	4.499447671	5.042599271	16.831348791	15.096629881	8.6438325781	8.6885543261	14.306382811	15.750759511	2.19002451	2.29777161			
BACH1_1	4.390780421	4.723868751	15.527418081	14.041675681	8.6833437391	8.7115875011	13.693410881	14.987107641	2.169475211	2.312198661			
BACH1_2	4.440212481	4.65170751	15.885240461	14.024801461	8.4863291781	8.1727818311	13.709056341	15.667824111	2.825626251	2.115016161			
BACH1_3	4.511256631	5.554337091	16.648355671	14.528117361	8.8550333311	9.1818806811	14.802006241	15.805287531	2.374812631	2.556130221			
BARD1_1	4.158176741	5.296281583	15.586838741	14.695503051	8.9915557271	8.4646649971	13.945986151	16.060625041	2.406075221	1.633115041			
BARD1_2	4.335650551	4.902047371	15.869758541	13.223698321	8.8210611991	8.8579900241	13.766995851	15.256500651	1.958774211	2.078947961			
BARD1_3	4.588784681	4.812632151	15.671023191	14.555034531	8.2102586611	8.4384315521	14.033171951	15.212773661	2.411088771	2.133138051			
BAZ1A_1	5.084745631	5.770744321	16.528887621	15.730359341	8.0610033391	8.3493560041	13.499161621	15.51047941	2.25512501	2.055144421			
BAZ1A_2	4.217926191	4.905352441	15.061449121	14.172669081	8.5625395421	8.7144943241	12.986466241	15.085984271	2.156143471	2.182690951			
BAZ1A_3	4.671175701	5.360578951	16.134521231	14.403307341	8.6043917081	8.9691347331	13.662772461	14.635682571	2.486089591	2.302118451			
BAZ1B_1	4.277872161	4.998059631	15.833282671	14.898576881	8.5313927931	8.7892786551	14.411633621	15.616872321	2.038654851	2.146007431			
BAZ1B_2	4.051678751	4.546443341	15.966278271	14.249891531	8.2548057591	8.5309798221	13.545361061	15.047952681	2.043480321	2.120836281			
BAZ1B_3	4.852099621	5.282260781	16.000313681	13.892612281	8.4819339281	8.259624031	14.504247471	15.509433961	2.117455731	1.893047851			
BCL11A_1	4.349884191	4.906357142	15.302636831	13.864437411	8.4076779371	8.7817911431	14.19751121	14.802070631	2.201153491	2.026630561			
BCL11A_2	4.351544671	5.214297751	15.840543311	14.340831911	8.5250804441	8.9472949441	14.085552061	15.577367591	2.291166031	2.163301621			
BCL11A_3	4.131842721	4.627802471	15.652389251	13.721669871	8.0634776891	8.4732897621	13.751424741	15.193236951	2.237938691	2.108439491			
BCL11B_1	4.25143051	5.100680431	16.380300231	14.157518731	8.3652787921	8.8165337111	14.007639721	15.81600491	2.350723131	1.901190851			
BCL11B_2	3.975273491	4.936745371	15.798253771	13.781351211	8.5262866281	8.7913196351	13.90183471	15.317332621	2.267019811	2.00245251			
BCL11B_3	4.341626861	5.111397461	16.686024411	14.314771561	8.475512491	8.894912221	14.211804151	15.787048931	2.049320191	2.136608151			
BCL7A_1	4.107457031	4.805028761	15.06831041	13.746363431	8.7506559951	9.0593982691	14.764270391	16.571601941	2.511678011	2.040449121			
BCL7A_2	3.69363221	4.185444461	14.660803671	13.708428671	8.2072938681	8.3765026721	14.495409041	15.656151491	2.193473111	2.316895471			
BCL7A_3	4.603150921	5.330725391	17.096057291	14.866202331	9.0444892781	9.0188316241	14.555926831	16.357780021	2.363931681	2.342610371			
BCL7B_1	3.591686581	4.932477731	15.526245311	15.310607841	7.8728733431	8.4602490851	14.137139671	15.614495681	2.057897561	2.111111111			

BCL7B_2	4.027381281	5.035612604	15.33530297	14.29849345	9.045624818	8.585545159	14.68163959	16.51347469	2.528724938	2.163717913
BCL7B_3	4.370284322	5.008553853	16.16021438	13.84336491	8.609372248	8.762053141	14.38687638	16.4579896	2.60386928	2.408037129
BCL7C_1	4.29554481	4.626629541	16.16002273	14.90667108	8.283354676	8.282570345	15.25629268	15.69676992	2.328542971	2.154382712
BCL7C_2	4.281210528	4.864199092	16.37649619	13.88449803	8.610972586	8.885769073	14.82956944	16.39825814	2.167817034	2.23782883
BCL7C_3	3.865374053	4.530397722	15.86809964	13.91142174	7.233194194	7.33733594	13.35406953	15.3579201	2.105877569	1.802677685
BCLAF1_1	4.348291579	5.26244574	16.87582275	14.0111886	7.678238863	8.37485172	13.22975604	14.7349268	2.115461244	2.154261342
BCLAF1_2	4.154378546	4.803617115	15.20506066	13.64210689	8.042578834	8.360330132	13.27981527	14.45006821	2.364476519	1.969189082
BCLAF1_3	4.411113843	5.671398827	16.38756139	14.04166366	7.414264062	7.509437189	12.94572299	14.03641293	2.106979541	1.962727186
BLM_1	4.359243697	4.998784736	16.21289132	12.94654119	9.906301837	9.317882731	15.77963999	16.09987745	1.951363539	2.597773337
BLM_2	4.21686747	4.930045984	15.3710714	13.77564422	8.685795533	9.160389796	14.71142229	15.96186695	2.278907604	2.076171999
BLM_3	4.617228978	5.247763978	15.71757457	15.79500387	8.81006282	8.967298073	14.5518336	16.24426378	2.265154874	2.072035613
BMI1_1	4.145066241	4.934603278	15.01071975	14.53049364	8.752421	8.903061678	14.07219343	15.79132812	2.60502478	2.200232707
BMI1_2	4.261339123	4.869981631	15.11899822	13.57218418	8.272073193	8.998690695	13.93791773	15.51102498	2.293008413	2.042163088
BMI1_3	4.076540008	4.704540193	15.3716823	14.02869012	7.978264435	8.414842146	14.47158123	15.3619366	2.227082387	2.0959299
BRCA1_1	4.601126658	5.178257666	16.26509247	15.15613018	8.966663596	9.462723663	14.60387243	15.84732727	2.176015474	2.244614469
BRCA1_2	4.336457395	4.918052456	15.67903467	13.92776419	8.461040872	8.776647057	13.2571626	15.49578448	2.109791303	2.294816911
BRCA1_3	4.661926803	5.226426868	16.86041284	14.2541528	9.159650641	10.19163015	14.56292448	16.06829671	2.126815469	2.337350071
BRCA2_1	5.002264751	6.056014795	17.99792915	15.47497547	9.639137806	9.763723434	14.95773422	17.1224908	2.607374511	2.503459326
BRCA2_2	4.963162001	5.518928466	16.8406353	14.6244508	8.365499403	8.617860096	14.5829683	15.73179945	2.037657983	2.218642454
BRCA2_3	5.030723486	5.895822109	17.21234224	15.23219673	9.577889328	9.740210787	15.06407343	16.71390577	2.809717865	2.707513972
BRCC3_1	4.56170711	5.216037369	15.90753988	14.65871692	8.904045864	8.903160811	14.95149137	16.02666786	2.525678796	2.52654262
BRCC3_2	4.407946038	5.281729515	16.42636513	14.58554502	7.959276043	7.811598374	13.77241165	14.64706327	2.368467448	2.114113696
BRCC3_3	4.494642951	4.984073196	16.585183277	14.09726147	8.455559544	8.612298243	14.33384705	15.7616143	2.033152033	2.029638253
BRD7_1	3.741891207	4.035584963	17.09714772	13.54934135	6.993591688	6.942379182	14.1494636	15.96082276	2.117913558	1.874884561
BRD7_2	4.100643426	4.709159051	15.60810523	14.66640942	7.253982897	7.526929517	13.39508248	14.69015617	2.047580661	1.631110529
BRD7_3	4.047079161	4.923029734	16.32754225	14.79860201	8.059766002	7.943140646	14.32108134	15.60084277	2.505222999	2.167664316
BRD8_1	4.390562786	5.294217477	17.52535959	15.20155866	7.920664649	8.111861424	13.86284856	15.54395589	2.043126514	1.980409393
BRD8_2	4.795585396	5.339893677	18.3717661	16.69237349	7.713532722	7.62399741	13.67024957	14.86297948	1.825662901	1.820683195
BRD8_3	4.419060634	4.704310527	16.18922528	15.09053862	8.027062065	8.243714215	13.31328751	14.4924462	2.096610079	1.689106133
BRD9_1	4.897458386	5.971810089	17.91577333	16.85777061	9.36024223	9.315822219	14.82847433	17.02498126	2.24022045	2.308676924
BRD9_2	4.819785007	4.802419256	17.41638601	15.67481471	7.59019667	8.545558498	14.77071999	16.2491176	1.850796616	2.137905488
BRD9_3	4.436633617	5.183173346	16.28753883	15.30425203	8.571805769	9.082784658	14.28102693	15.76757524	2.606485245	2.162229205
BRIP1_1	4.044726346	4.599655818	13.45764436	12.73737252	8.448811255	8.704609279	13.66311939	15.23420069	2.256737028	2.195396466
BRIP1_2	4.186516429	4.552701506	14.74875998	13.54388011	8.657850566	8.309657883	13.9296325	15.00145037	1.93698535	1.972272053
BRIP1_3	3.931272836	4.477209813	14.02073627	13.08363366	7.865685261	8.400345508	13.96675237	15.81167572	2.260687057	2.423049144
BUB1_1	4.24622522	5.002623048	15.76684875	14.29897202	8.298332201	8.797292334	14.09323458	15.0976783	2.136305478	2.379755584
BUB1_2	4.079353951	4.917219775	16.10139616	14.71680796	8.335144319	8.881921927	14.46609693	15.73695521	2.176800066	2.311283687
BUB1_3	4.246226223	4.930627511	14.27054204	14.69079335	8.655763861	8.417659954	14.11693016	15.15706266	2.241009712	2.113896467
CABIN1_1	4.761279553	5.251321264	16.86441018	14.2046925	8.306023029	8.66446846	14.01955504	15.61282733	2.248953975	2.110885831
CABIN1_2	4.62576741	5.14946854	16.86116319	13.42243367	8.059930723	8.111810405	13.24554961	14.48496429	2.350188951	1.903141128
CABIN1_3	5.056616821	5.20155344	16.16363087	13.7548668	8.096710445	8.01100464	13.5595573	15.57935673	2.006677768	2.140211462
CBX2_1	4.267686751	4.791108731	15.32716385	14.54360354	7.906687872	7.96812749	13.64419493	15.44345723	2.125399192	2.187301131
CBX2_2	4.209428398	4.865633645	16.76240314	14.97025501	7.769971638	7.870023103	13.753211437	15.71328351	2.474793767	1.947891316
CBX2_3	4.351305392	4.997116366	16.51886721	13.8747993	8.544537713	8.538667968	14.63961429	15.19177634	2.199575445	2.107828466
CBX4_1	4.266329985	4.690394511	15.86716942	14.46495899	8.267397611	8.113060803	13.79074942	15.10523043	2.191371795	2.132334705
CBX4_2	4.544633721	5.183178556	15.95674294	14.51162946	9.162224069	9.437371366	14.45899728	15.45873643	2.028259736	1.931096247

CBX4_3	4.5385042461	6.00035823	14.84496755	13.03950104	8.07052752	8.31095767	13.84293765	14.65356316	2.15449744	1.99078888
CBX6_1	4.5620766311	5.14088752	16.56845069	14.66770236	8.57775113	9.03455955	14.8778049	15.51996527	2.20881215	1.86201333
CBX6_2	4.2739162111	4.89424575	15.98906293	14.30338463	8.93115054	9.18234912	14.51489621	16.21082209	2.26482549	2.13426049
CBX6_3	4.0438835161	5.24643577	15.96614911	14.59685514	9.10475896	9.28758525	14.68817406	15.8194789	2.36777	2.11671287
CBX7_1	3.7775683321	4.79852075	15.93051726	14.26193488	8.30208166	8.00809740	14.53051151	15.64056053	2.50336587	2.49983553
CBX7_2	4.6864516131	5.27980321	17.36101856	14.90668365	8.72601966	9.46220546	14.73386326	16.58692478	2.34835465	2.23444202
CBX7_3	4.3203422051	4.73584548	16.33735667	15.00308417	8.55776459	8.76157503	13.5746259	15.41603956	2.13834035	2.27603824
CBX8_1	4.311269294	5.25581939	16.64943123	14.46296836	9.66019196	9.69143801	15.05114952	16.04461039	2.11774704	2.35961501
CBX8_2	4.8778722691	4.96449862	16.4097267	14.45120263	8.87516306	9.18372696	14.49758473	16.14099415	2.27054466	1.93601287
CBX8_3	3.6866250141	4.77345347	16.19063935	14.43033205	8.68805789	8.44765044	15.05655804	16.11838617	2.39270754	2.05491416
CCDC58_1	4.4985399981	5.03563905	15.64650234	14.60175503	8.36613809	8.25696187	13.8637879	15.71941598	2.15421898	2.16385272
CCDC58_2	4.1457909941	4.80635601	16.08605513	14.30612198	8.16635259	8.84008236	14.3952436	16.27476862	2.23605136	2.06506064
CCDC58_3	4.6330097051	5.48010321	16.94831014	13.98985998	8.97990608	8.97910567	15.01843781	15.94062514	2.11628901	2.02852981
CCNA1_1	3.7948243041	4.81284791	14.99982605	13.16726193	8.21461603	8.00606484	13.70439974	15.05979625	2.11416113	2.17001051
CCNA1_2	4.3203207531	5.12380662	16.42447258	13.43226833	8.06372233	8.48700158	13.56792744	14.34024906	2.05384079	2.24208035
CCNA1_3	4.6470707661	5.09686828	15.76500519	14.12502113	8.60153477	9.07704973	14.18295519	15.65694765	2.22727961	2.15956629
CCNE1_1	3.7926607471	4.74130512	15.02005551	13.91496864	8.12643627	8.55835316	13.7531088	15.75698313	1.95818739	2.09379458
CCNE1_2	4.2202325271	4.89130434	16.40434895	14.55263467	8.27943726	9.26267736	14.9173498	16.55673398	1.78210929	2.10506606
CCNE1_3	4.1499272831	5.32432720	15.60281548	14.51721263	8.45881274	9.07263602	13.9983383	15.62646608	2.20906584	2.4029364
CCNH_1	7.7179535121	8.02322241	24.22675286	22.35279695	9.28294375	9.80582655	15.76167719	16.17170341	2.63845558	2.65324243
CCNH_2	5.1203362161	6.02734830	19.59722202	16.10484942	8.35475332	8.49664119	14.0899024	15.25809753	2.27925403	1.83652086
CCNH_3	7.3260073261	7.94635131	23.14019919	21.05581586	9.43482029	9.91536811	15.4133514	16.77625357	2.27057234	2.48060517
CDC40_1	5.5537080151	6.80583489	15.95159516	15.27864746	8.02495806	8.92163024	15.28719429	16.17638372	2.84559113	2.17061114
CDC40_2	4.466380631	5.35639413	14.64773922	16.33891009	7.81574963	8.49433106	13.79677183	13.90289201	2.10609496	1.32320119
CDC40_3	3.7421190091	5.77800304	16.92550079	16.1776368	7.98277343	8.55470751	13.84796818	14.9186272	2.35166668	2.58224266
CDK1_1	4.4954316941	5.32776191	17.01048951	11.86058055	8.7255573	9.16183339	14.26783323	16.27433275	2.67414791	2.78464863
CDK1_2	4.0582423971	4.68439775	16.64652519	13.36549872	8.64522016	9.01683321	13.80666441	14.86539182	2.52523072	2.26248666
CDK1_3	4.3063889671	4.84007923	14.51062215	14.03710166	8.37432058	8.22428012	14.41226856	16.56323841	2.20296744	2.76173663
CDK7_1	5.2958895091	5.71436204	19.66262031	16.16957769	8.60277136	9.41268231	13.5680943	15.92626871	2.32544599	2.53911957
CDK7_2	5.9498540731	6.25347391	20.67786071	18.09404106	8.31408001	9.08392651	14.4707219	15.62332161	1.96154102	2.19196314
CDK7_3	7.3108808291	7.57581038	23.15421031	20.55909816	8.79016564	9.16582415	15.54647763	16.63033279	2.42220782	2.42856446
CDYL_1_1	4.5512699461	5.32801868	16.02634047	14.54133684	8.30875471	8.19772346	14.0123196	13.99664205	2.10844892	2.45161053
CDYL_1_2	4.6843960271	5.86606667	17.30271997	15.53452547	8.72950575	8.94663754	14.6453662	15.43281595	2.24411453	2.19542601
CDYL_1_3	3.9428856191	4.97123249	16.82225176	15.20960361	7.56316610	7.83142453	14.2084926	15.61565771	2.26024563	2.17259637
CDYL_2_1	4.4234789221	4.93353843	15.81486043	15.20595873	8.54056396	8.81932322	14.5042231	15.73444079	2.24354866	2.28640192
CDYL_2_2	4.2156352091	5.03006789	16.12194212	14.73545466	8.53245326	8.25625394	13.9721806	15.46562575	2.34074400	2.09415312
CDYL_2_3	4.4254312271	4.72467783	15.08636101	14.20680424	8.18323149	8.57003779	13.5552079	15.66951813	2.34603484	2.20756285
CETN2_1_1	3.8134073611	4.35439950	15.95534364	13.87427241	7.80227647	7.92835914	14.1418184	15.25990306	2.17365302	1.98733347
CETN2_1_2	3.522635951	4.33517425	15.75680277	12.84529361	7.95603354	8.03498723	13.9187777	15.1087163	1.98577355	2.26149673
CETN2_1_3	3.8626230521	4.81474230	15.92210685	14.77282299	8.47914276	7.73438878	14.0122237	14.92624373	2.34210139	2.25170365
CETN2_2_1	4.0164959781	4.53119681	15.60462671	13.28714466	8.42420923	8.62271803	13.3602811	15.30532711	2.18472796	2.04469666
CETN2_2_2	4.6387080831	5.15024639	17.19579135	14.62031977	8.21935223	8.42750026	14.2515608	15.40986005	2.14638742	2.24941943
CETN2_2_3	4.4024474741	5.18084503	15.54640659	13.64496529	8.41448278	8.14632541	14.011622	15.28386021	2.42756083	2.23936023
CHAF1A_1	4.9667419381	6.19639853	20.81202541	17.22080799	8.93927358	6.74436852	12.1237950	13.33694393	1.72024403	2.12181566
CHAF1A_2	3.9437274691	4.38165724	17.06965844	14.62723942	7.37968464	7.00659340	13.2395855	14.6804309	1.88735164	1.90434056
CHAF1A_3	3.9613505871	4.47859515	17.60982953	13.88663081	6.95362303	7.30427764	13.1358284	15.0542576	2.18396784	2.03386842

CHAF1B_1	4.450008536	4.376699454	20.16054169	16.55802721	4.576455295	4.656522659	10.00320374	10.12658228	1.903122859	1.760051671
CHAF1B_2	5.313857276	6.26331138	24.30070942	19.74168056	4.532017531	4.458692145	9.169380593	9.862656396	1.742401742	1.730444569
CHAF1B_3	4.927833965	5.866064539	24.30012312	19.71424248	4.266410652	4.262633207	8.857794859	9.562486469	1.966457604	1.696107823
CHD1L_1	4.353250164	5.079557713	17.36869697	14.09768792	8.496727181	9.018347848	14.5829286	15.68707079	2.348321122	2.609843409
CHD1L_2	3.822125363	4.502940316	14.47812975	13.5313245	8.492629192	8.463799963	14.14888526	15.49395049	2.156755774	2.07969519
CHD1L_3	4.422828395	4.822245946	15.80886162	13.74692908	8.092633261	8.616959229	13.69222666	15.26169477	2.135807925	2.310209208
CHD2_1_1	4.524372929	5.041723944	15.70551475	14.29023302	9.919439028	10.17365296	15.98225416	18.14029014	2.717003063	2.139056658
CHD2_1_2	4.023492431	4.474793561	15.27793868	13.393731641	9.34558064	9.302662037	15.73151075	16.95954737	2.396109341	2.262370143
CHD2_1_3	4.131315861	4.282368954	14.87169459	12.75697254	9.333381197	9.831529371	15.93116743	17.44090562	2.086800387	2.49687002
CHD2_2_1	4.487684376	4.890061589	16.31132437	14.80512326	8.389672226	8.87506431	13.94849765	15.63600996	2.023345361	2.047465731
CHD2_2_2	4.080245798	4.986502663	15.29894336	13.67907943	7.720770973	8.242356541	12.90282954	13.71751632	1.852213891	1.884256345
CHD2_2_3	4.456651446	5.115370605	16.10881634	14.71246207	7.914658386	8.130791477	12.2702032	13.86729006	2.126745509	1.931685285
CHD3_1	4.261684436	4.906066635	15.36224281	13.50207718	9.101851265	9.16369611	13.9184011	16.31447929	2.517874079	2.510233403
CHD3_2	3.961705377	4.760741904	15.33475575	13.55489043	8.960393019	8.485103695	15.506532	16.37532796	2.771549638	2.15339051
CHD3_3	4.649463824	5.461332342	15.73311153	13.75968992	8.671854172	8.756119921	14.5098646	15.51549744	2.064338552	2.582290944
CHD4_1	4.85218147	5.580419261	16.71159127	16.37634191	8.166804798	8.246674358	14.10193529	14.96358068	2.044325127	1.901823333
CHD4_2	4.493398688	5.739679094	16.27353816	15.00257857	8.199263639	8.837414759	13.6354616	14.63546477	2.089691716	2.012700295
CHD4_3	4.681177392	5.045736847	17.03341336	15.03052612	8.517115411	8.690503748	14.70297089	15.5595165	2.239336817	2.165185281
CHD5_1	3.920357868	5.085556113	16.34879126	15.01142965	7.942236718	8.196243263	13.69876343	14.73770979	2.238471643	2.151112765
CHD5_2	4.059405941	4.791809904	15.46784738	13.63011816	8.155671863	8.151066288	13.85648335	14.76310009	2.012478155	2.387798166
CHD5_3	5.056369333	5.877811181	17.79275554	16.56108732	7.142300515	6.894590306	12.1788743	13.16084492	1.874726842	1.808638828
CHD6_1	4.332719389	5.045210166	16.21345231	14.99768069	8.380468378	8.474072017	14.69825112	15.95805249	2.536256324	2.412700584
CHD6_2	4.575403623	5.394612249	16.63409182	15.46305451	8.20559668	8.481447964	14.20088884	15.50097628	1.900410719	2.144841396
CHD6_3	3.918552366	4.506531907	15.76061524	13.32058535	7.537986125	7.729768766	13.036911117	14.89725345	2.056739802	1.959490612
CHEK1_1	5.416585964	5.773116521	16.09686011	15.08028441	8.723450948	8.753450436	14.1194588	15.65181747	2.447685871	2.213083845
CHEK1_2	5.031951936	5.742930567	16.4148823	14.77071972	8.652653507	8.973600219	14.19580439	15.61486698	2.367020334	2.118097386
CHEK1_3	6.779366647	6.136051864	13.52021631	16.66645118	9.15984126	9.339674403	14.99187882	16.45065465	2.502931543	2.168340413
CHEK2_1	4.481273633	5.126839517	16.53366335	15.06271675	8.207030342	8.791125166	14.1949651	16.18444172	2.105449267	2.201321958
CHEK2_2	4.454648456	4.866982496	16.24607146	14.30002691	8.621919159	8.846500019	13.68063329	15.32229321	2.172439394	2.016393277
CHEK2_3	3.907015824	5.112267101	15.66075325	14.0833583	8.010400867	8.201292006	13.48724479	14.77558042	2.150214453	2.037684895
CHFR_1	4.753514984	5.107610906	16.38395755	15.72292243	8.570022586	9.364104779	14.35269695	15.55027949	2.320421558	1.899193906
CHFR_2	4.356698185	5.057995895	16.16657366	14.76075947	8.605437915	8.785042377	13.97017317	16.1901283	2.28456332	2.106476772
CHFR_3	4.450321116	4.713907391	15.3863216	14.69108177	8.807844944	8.6531179	14.49374251	15.30635749	1.996785734	2.194570136
CHTF18_1	4.236428209	5.121814706	14.79181694	13.20302858	7.564573622	8.171619163	13.30345736	14.62517353	2.349616389	2.082653363
CHTF18_2	5.075899046	4.786031321	16.45303725	13.30940633	7.803494268	8.650053941	13.3987453	14.02859331	2.096724255	2.122851072
CHTF18_3	3.600275273	4.315056372	15.58444715	13.22153921	6.328879065	7.109370691	12.90414336	14.45026697	2.247488102	2.268104478
CLK2_1	4.351741568	4.78753855	16.73236541	14.0570942	8.616732648	9.12634448	14.33217659	15.81113742	2.067499336	1.996032156
CLK2_2	4.218499723	5.034398094	15.29830375	13.81834292	8.623053812	8.845544682	14.7279642	16.17195773	2.326994066	2.172976831
CLK2_3	4.193622716	5.001531663	15.44813164	13.95897934	8.666529478	8.842555732	14.79446633	15.4933147	2.211302211	2.214568372
CLOCK_1	3.243974632	4.851575475	13.4848699	12.2028017	7.791565403	8.654561108	12.99981315	14.32659102	1.990675105	1.826761353
CLOCK_2	4.350060351	5.02077329	15.86688878	13.71399843	8.561900449	8.708557744	14.09370317	15.58754724	2.247956903	2.339520407
CLOCK_3	4.371062283	5.30993529	17.36982622	15.05839167	9.169305464	9.041880533	15.12709059	16.26631933	2.396588496	2.351210224
CLSPN_1	4.774676645	5.153895332	16.55221407	14.55694796	8.361958486	8.94702	13.84055442	14.67171436	2.337732884	2.015137407
CLSPN_2	4.204219075	5.108785916	15.72834977	13.74742601	8.035494208	9.427354131	12.87622984	14.57348027	2.243963364	2.556733512
CLSPN_3	4.269749096	4.748174164	15.76656428	13.70161226	7.349745033	7.398033444	14.07682142	15.16850045	2.289719279	2.320129861
CREBBP_1	4.169607471	4.775795576	14.73979108	13.16556561	9.962037285	10.28097231	17.05934004	18.23929367	2.919040132	2.663326386

CREBBP_2	4.359178113	4.880711907	15.151402231	13.17378143	9.850145662	9.842164685	16.08984246	17.92878938	2.816610276	2.498415336				
CREBBP_3	3.807191801	4.568812904	15.013233355	12.77952956	9.524966168	9.184505607	16.05663278	17.48308259	2.759377886	2.787123711				
DCLRE1A_1	4.589737334	5.036181967	16.98957664	13.84466366	8.986247363	9.569236953	14.2034048	15.46731331	2.297773035	2.117031017				
DCLRE1A_2	4.177435116	4.820104659	15.75431232	14.47478156	7.734616142	7.856041131	13.09549281	14.97556919	2.212513364	2.256825483				
DCLRE1A_3	4.529739232	4.773496263	15.25416844	14.5433844	8.026204864	8.245405304	13.9093998	15.46319431	2.126782812	2.059762647				
DCLRE1B_1	4.486321796	5.051402719	16.78414952	14.91014361	8.398339342	8.344624649	14.01078226	15.64741773	2.056825087	2.239571184				
DCLRE1B_2	4.20362998	5.219146482	15.85004933	13.36130343	8.827684605	8.336092504	14.48645105	16.59702622	2.215169921	2.362300729				
DCLRE1B_3	4.202980196	5.085435313	16.00380276	14.2164613	8.760527762	8.721226478	13.85005914	14.93064465	2.248444329	2.159617703				
DCLRE1C_1	3.608480033	4.196927172	15.51365005	14.55239798	6.821999794	7.609560978	13.89742194	15.19009715	2.368320123	2.158412491				
DCLRE1C_2	4.14546782	5.070422535	15.37586611	14.28609914	8.429554381	8.777053175	15.03257964	15.62025176	2.465947432	2.452166263				
DCLRE1C_3	4.493035951	5.224207324	15.73147295	13.66378967	8.490186009	8.377417686	14.23253894	15.4094494	2.234458283	2.157390763				
DDB1_1_1	4.031007752	5.658713979	17.14239046	16.24706732	5.05119013	5.535113749	9.584175768	10.00281842	1.562730355	1.154603105				
DDB1_1_2	4.261203231	4.778230434	16.73104952	14.86411316	8.709623493	9.176169465	14.17117971	15.67136347	2.281616237	1.972384348				
DDB1_1_3	5.014584609	5.686776651	17.03544173	15.24034689	7.326026066	7.69057354	13.35657771	14.34318134	2.216350984	2.402452355				
DDB1_2_1	4.344166891	4.685458069	14.91821413	14.84419323	8.866735869	9.066464628	14.32445579	15.35096039	2.54735096	2.305242226				
DDB1_2_2	4.296651976	4.952211355	15.51627926	13.53658608	8.701600959	9.104396782	13.69075736	15.64761563	2.016918544	2.315605862				
DDB1_2_3	4.022754426	4.947436101	16.02198315	14.4913665	8.14094316	8.918138252	14.3705535	15.51743646	2.294624	2.030936656				
DDB2_1	4.361622424	5.333434778	16.02344432	14.73633126	8.832571277	8.759616682	13.942482	15.80623204	2.200213966	2.033683971				
DDB2_2	4.469170875	5.102437669	15.45007664	15.19752544	9.154699883	9.495449398	14.56349674	15.17734396	2.591377337	2.595928618				
DDB2_3	4.157340107	4.668860316	14.69541843	13.18473827	8.516631375	9.110639127	13.96185705	15.16600589	2.259708337	1.968844151				
DDX1_1	4.160171856	4.559895194	15.39284741	13.74089901	7.773478395	7.841035896	13.21610891	14.24432736	2.053370811	1.908830933				
DDX1_2	4.31834983	4.233056882	15.19688145	13.52751157	7.669216087	7.551338705	12.96901823	14.0198121	2.046099588	2.124020524				
DDX1_3	4.499536608	5.371975701	15.64151092	14.1423239	7.559414147	7.916800439	13.03844205	14.03372835	2.096828547	1.836746254				
DDX11_1	3.690664379	4.851397189	14.59747163	13.65336847	8.223208117	8.519039169	13.86686774	15.54663271	2.099431208	2.050709981				
DDX11_2	4.279287451	4.685637161	16.05304484	15.24716858	8.566877511	8.390322317	14.23034504	15.58547949	2.126816783	2.05859193				
DDX11_3	4.534457192	4.914408266	16.04686704	14.00707651	8.417842365	8.409250175	13.98408844	15.10158093	1.997884488	2.064916216				
DDX3X_1	4.341189938	4.742790551	16.75109993	14.69577009	8.244472615	8.088823688	13.90813173	15.00075796	2.369289561	1.919734351				
DDX3X_2	4.556137165	5.731919608	19.30022871	16.99960499	6.582210465	7.366352408	11.03463661	12.77618515	1.679833333	1.849294879				
DDX3X_3	4.667054376	5.266894322	15.53909465	15.71100008	7.487386716	7.73053255	11.81728719	13.57055482	1.782310205	1.881668673				
DDX5_1	4.779680525	5.618847367	17.40033449	15.66253966	8.397883416	8.306896521	13.90330885	15.71985588	2.251320549	2.411852235				
DDX5_2	4.986025855	5.975952303	16.6352729	15.26572152	8.877520086	9.261804496	13.96625864	14.93595576	1.82602197	1.938522263				
DDX5_3	4.893323275	5.544654034	18.31488232	16.02701367	8.36307247	8.006079488	14.26737	15.24002938	2.177501086	2.237499712				
DDX58_1	4.477115827	5.101555439	15.46437496	13.17687726	8.558580382	8.526139646	13.95985277	15.36042543	2.262321792	2.105052876				
DDX58_2	4.266244495	4.826219857	14.95911584	13.77468096	8.761055955	8.526680446	13.92682456	14.94698234	2.276213942	2.239468937				
DDX58_3	4.319560074	5.086604026	16.81514039	14.00569851	8.459858932	8.511695507	14.32854227	15.34278619	2.221645036	2.310617564				
DHX15_1	4.732425368	4.500191498	15.45089883	15.26733792	5.430428934	6.34525365	10.19830455	11.85699588	1.152801358	1.272332786				
DHX15_2	5.334938557	6.226351635	18.2174677	14.94474523	7.791155294	8.336545894	12.68954333	14.38684326	1.537287095	1.535427406				
DHX15_3	5.382200537	4.802966014	16.70700609	14.74650107	7.416580369	8.254654549	12.87648868	14.29935391	1.665021292	1.714426043				
DHX36_1	5.653462804	6.305361762	21.25399886	17.25490862	8.547021017	7.511797433	13.005687	14.78220304	1.738647716	1.935806617				
DHX36_2	5.405419715	6.259874769	19.8748641	16.97823453	7.960306643	7.834966303	12.73787557	13.81852642	2.111445063	1.972325278				
DHX36_3	4.975927032	5.878211321	18.95662948	17.53763021	7.293574528	6.815922186	13.72037892	13.87857471	1.953365323	1.835229763				
DHX8_1	4.815044431	5.959055939	17.14686623	13.96895767	8.311750921	8.393370657	14.21565429	15.3167776	2.329564919	2.338718192				
DHX8_2	3.755605381	5.435569401	15.46172741	14.75956077	8.727600226	8.786747044	14.26867885	15.61189019	2.261026668	2.134339605				
DHX8_3	4.850444624	4.266938751	15.50863908	15.37444206	7.102278491	7.582862196	12.75682048	14.07469171	1.524057204	1.92818836				
DHX9_1	5.329365148	6.693766936	19.8497115	17.18158001	8.437478632	9.065567219	13.90707896	15.41455755	2.191264418	1.868385053				
DHX9_2	5.058442845	5.454761056	16.95163911	14.80275767	9.119412151	9.331190968	14.67867969	16.79900512	2.288247353	2.422700427				

DHX9_3	4.717467358	5.296101324	16.59873197	15.0594B166	8.632246099	8.715666333	13.49729372	15.34787008	2.086735432	2.090680057				
DICER1_1	4.781870147	5.409872487	17.96555154	16.74730108	8.669767191	8.601218834	13.81911192	15.53860608	2.081937266	2.129555697				
DICER1_2	4.762602043	5.820535917	17.50408345	15.68175341	8.999791455	8.807327632	14.46206431	15.80756675	2.254181331	2.095871521				
DICER1_3	4.601553024	5.567105748	17.006131 61	16.55150998	8.695631868	8.820433329	14.17888609	15.64721258	2.196771361	2.212377572				
DMAP1_1	5.880198286	6.736642997	21.85007412	20.43705604	7.61097622	7.633001141	13.32990372	14.51714168	2.527407843	2.262044653				
DMAP1_2	4.6773190	5.807232895	22.27105336	20.12453552	7.271356666	7.367307971	14.59315494	14.99014803	2.356290201	2.350539386				
DMAP1_3	7.28983889	7.931162227	24.81112906	21.98753654	8.152752875	9.030516503	14.37919077	16.01771442	2.340702211	2.073374906				
DMC1_1	4.145764094	4.936416887	15.58730044	15.70842176	8.456805019	8.680639226	13.41839144	15.61939166	2.019570653	2.165477669				
DMC1_2	4.505136206	5.188540369	18.0402961	15.16088609	9.086623386	8.599076314	14.57446041	15.59340067	2.295485171	2.016707761				
DMC1_3	4.434810042	4.964899658	13.34603141	13.39137869	9 184843 62	14.27422674	15.99544302	2.366655873	2.202531881					
DNA2_1	4.911077531	5.639169366	15.87323317	15.24948481	8.083388372	8.327402135	13.52923171	14.78046169	2.429721712	2.19764084				
DNA2_2	4.425322668	5.125307619	16.73884682	14.94506516	7.415066779	7.938769167	13.29620791	13.97052666	2.762058744	2.332790334				
DNA2_3	3.967241727	4.052405692	16.11978962	14.29450874	6.004336627	6.792943428	11.85773024	13.42744401	2.293181916	2.320454545				
DNMT1_1	3.704572597	4.346018624	13.13377256	13.0467677	9.465480768	9.712019126	16.36168444	17.0844701	2.664072377	2.417812468				
DNMT1_2	3.895002181	4.458304114	16.86553706	13.05202832	8.744854385	8.749890204	14.27613655	15.14609422	2.350164912	2.280328367				
DNMT1_3	4.363863353	4.923044671	15.21559354	14.26103174	8.549294669	8.960326604	14.28969359	15.57692688	2.153755253	2.066717654				
DNMT3A_1	4.241034533	5.31469826	16.08198772	14.97818024	8.403088301	8.913463056	14.32288475	15.89761763	2.131201247	1.963669942				
DNMT3A_2	5.15734642	5.19510527	17.17499961	15.158762	9.042261371	9.094933646	14.78629032	16.25126305	2.512811155	1.953912406				
DNMT3A_3	3.877607918	5.103568573	16.06419662	14.93322004	8.602880705	9.038142072	14.44860497	15.05550263	2.065382843	1.934129324				
DNMT3B_1_1	4.032957502	4.659834217	16.15469562	14.46469952	8.604443113	8.662583866	14.74413878	15.78747497	2.257215982	2.204834057				
DNMT3B_1_2	4.73393469	4.69419916	15.86157164	15.14330832	8.563501 78	8.941317074	14.41957764	16.07019327	2.307754377	2.255481904				
DNMT3B_1_3	4.380602844	5.111141498	16.06133201	14.44413106	8.411886076	8.79835375	14.36466636	15.53484986	2.193426273	2.187603791				
DNMT3B_2_1	4.307140294	4.490198518	15.53057024	14.03501997	8.465059948	8.959468529	13.72650551	15.63274895	2.052418247	2.22664932				
DNMT3B_2_2	4.235609666	5.124869018	16.21137827	13.83124137	8.090632903	8.006207867	14.26668572	15.04255319	2.242366752	2.173593274				
DNMT3B_2_3	4.093902425	4.514190845	15.13160648	13.82500767	7.990230683	7.746005921	13.86433273	15.30009031	2.301608823	1.853782876				
DNTT_1	5.392689106	5.017126738	14.16100873	16.97033898	10.65375303	8.053515215	15.73886666	17.21735569	1.885902876	3.310294653				
DNTT_2	4.716488111	5.262613088	15.94175723	14.12311992	9.518206002	9.72959955	14.06334683	15.71096752	2.209702724	2.517484731				
DNTT_3	4.597388955	4.710615381	16.3449426	13.96172008	8.423709745	8.647322469	14.33497112	15.66953309	2.55452511	2.138234818				
DPF2_1_1	4.420725166	4.708941113	16.50316633	14.76705466	8.682796661	8.706616 12	14.4318597	15.43133309	2.122747144	2.197907946				
DPF2_1_2	4.155852073	5.034086448	15.70560682	13.96282665	8.152892876	8.416405211	13.71908157	14.64095394	1.947095536	2.175749662				
DPF2_1_3	4.450359092	4.785782829	16.06661803	14.85277307	7.341035972	7.399486741	12.06236557	13.56561082	2.172315715	1.596619551				
DPF2_2_1	4.366217527	4.808805776	15.43867387	14.34390735	8.449616097	8.144752164	14.18880045	15.11127698	2.4160596 15	1.935247681				
DPF2_2_2	4.206899676	4.949699645	16.21385703	14.67508881	8.222217897	8.236611873	13.83633113	15.64750153	2.205613899	2.318666929				
DPF2_2_3	4.093305722	5.007141724	16.56234235	14.23205121	8.239270146	8.70472956	14.59537078	16.45233509	2.304299526	2.115305509				
DROSHA_1	5.332976186	6.084401396	17.97469105	16.31082912	8.320097902	8.932213747	13.65527146	14.71849223	2.067681483	2.004490412				
DROSHA_2	5.913770806	6.705413597	19.58343459	17.45234875	8.424793317	9.052708 15	13.96771105	15.73070044	2.130062838	1.941896577				
DROSHA_3	4.983131653	6.444940651	17.60925951	17.3669597 8	8.467389574	8.827316765	14.42333936	15.21450577	2.300363367	2.197282422				
DTL_1	5.630603706	6.471547931	16.91630921	16.16341643	9.049249138	9.794319295	15.02753508	16.02105191	2.229251608	2.261201321				
DTL_2	4.657964578	5.320286658	16.37179703	14.30634282	8.326445709	8.637612878	14.10053968	15.36030079	2.001161159	2.105868815				
DTL_3	4.746009589	5.649348152	16.59082136	14.23716814	8.275254125	8.340828602	12.78607331	14.52624084	2.274187073	1.737409261				
DUT_1	5.412950228	5.154353745	16.77804635	12.87219423	8.545047897	8.896284722	13.77169129	14.86331377	2.191994243	2.106717413				
DUT_2	5.408151744	5.356804652	15.50661798	12.48575028	8.273217507	8.601305424	13.42434896	14.91001686	2.149692833	2.212145368				
DUT_3	4.273761966	5.570057072	15.74419272	14.24313725	8.924703922	9.660125134	15.2484818	16.78081943	2.408670141	2.528654462				
E2F1_1	3.790617835	4.381173967	15.96204595	13.06365115	8.526771149	9.097551541	13.86497446	15.88595626	2.116207148	2.24772002				
E2F1_2	4.381446976	5.376283431	16.77790865	15.2644354	8.694395482	9.206467459	13.81727745	15.80494637	2.401603965	2.303567752				
E2F1_3	4.128366628	4.532515066	16.20017291	14.39234401	8.169415261	8.354307359	14.17028671	15.47195766	2.069615159	2.184335936				

EED_1	4.578275014	5.282833226	15.48400673	14.00579151	8.845211963	8.49808203	14.91890347	16.43532413	2.168630083	2.003525863
EED_2	4.363012229	4.803386361	15.56606661	14.54515395	7.958305231	8.708235588	13.88255048	15.38491042	2.362344862	2.032844452
EED_3	4.506534984	5.290785498	16.37240282	14.5565883	7.796254369	8.301807205	13.34169679	14.28153897	1.845715238	2.021290539
ELP3_1	5.532568246	5.874981559	18.91621354	16.64606157	9.200713321	9.287627766	15.01360672	16.34664078	2.059427008	2.256400674
ELP3_2	5.376985202	6.065429116	19.53821026	16.40565034	9.436398336	9.830437677	16.44110795	17.74096901	2.625006586	2.307925419
ELP3_3	4.880018759	5.387825981	17.99832496	16.29301818	9.374378531	9.903872562	15.41789342	17.20291898	2.752103588	2.406125903
EME1_1	4.664021611	5.004486871	16.65080129	14.2676415	8.684188309	8.793513649	14.21393427	15.43646089	2.096062754	2.659739388
EME1_2	4.928585796	5.374492992	17.67416704	15.74723441	8.48433792	9.285157121	13.89555569	15.77586061	2.436775371	2.472711783
EME1_3	4.416282943	4.869566251	15.87424477	14.37089879	8.827073559	8.990039627	14.7172801	16.20284558	2.128442616	2.046938115
EME2_1	4.001559347	4.559676463	14.88012586	12.86436443	7.718635404	8.009642747	13.80020877	15.21239885	2.1525882	2.004145821
EME2_2	5.440103142	5.851520646	18.04186415	15.79392147	9.18024896	8.771668948	15.30955565	16.55698522	2.210432714	2.295283019
EME2_3	4.407822215	4.941600941	15.26190416	14.04301087	8.368810501	8.812126164	13.44878181	15.01531792	2.178024809	2.112663045
ENDOV_1	4.401554446	4.974766412	15.10327097	13.76447235	8.050255474	8.460590237	14.2903504	15.74651909	2.430016923	2.401737632
ENDOV_2	4.379976637	4.948352393	16.36549022	14.45408092	8.566140783	8.520336023	14.13273251	15.93302604	2.401674008	2.016735243
ENDOV_3	4.416660877	5.073173919	15.12911128	13.47154557	8.764861591	8.755245956	13.42886293	15.40575101	2.324300326	2.045556984
EP300_1	4.190176004	5.093535562	15.37441526	14.01720869	9.008058951	9.053421539	14.32475248	15.62944362	2.138473341	2.119339409
EP300_2	4.861144591	4.730887699	16.67855245	15.53811594	8.772170219	9.324179955	14.60561393	14.83545409	2.227151052	1.910252339
EP300_3	4.187151754	5.114787373	15.2934876	13.8362986	8.632093041	8.85357667	14.0661892	15.70817007	2.366282596	2.108252941
EP400_1	4.113286149	4.688812364	15.68636546	14.33905519	8.502896538	9.066879365	14.03641432	15.5792322	2.036429308	1.946760291
EP400_2	4.357958741	5.659070006	19.87325503	17.65452635	8.543485934	8.853946372	14.56558416	15.84305921	2.250971988	2.079004341
EP400_3	5.903050133	6.362870304	23.41412359	20.68165615	8.267620666	9.080553958	14.63443126	15.39414272	2.424151037	2.143615475
EPC1_1	4.160597639	4.476222478	16.46852056	14.35980451	8.683371692	9.070711044	14.31209707	15.52285155	2.224576865	2.095261193
EPC1_2	3.529651649	4.313504757	15.54379211	13.63188477	7.892118735	7.93201869	13.50613622	14.83573586	2.090121492	1.970264086
EPC1_3	3.516866018	4.113678112	13.44010309	11.82772757	7.311155968	7.336298147	11.54998096	13.69822342	1.801251717	1.870046751
ERCC1_1	3.946798227	4.214187836	15.44533607	13.11953819	7.438752476	7.258483164	13.40513757	14.37705689	2.010815318	1.965348556
ERCC1_2	3.847291126	4.165344389	15.56214061	13.73729122	7.57111468	7.951992641	13.77711208	14.64075967	2.074705644	2.203182375
ERCC1_3	4.459287532	4.897539204	14.94291709	13.02074724	8.5875174	8.557067337	13.94759311	15.39061764	2.389626745	2.411389318
ERCC2_1	4.277373656	4.833689739	17.42519925	15.26483566	8.707283379	8.662512285	16.58418485	17.96956144	2.61812186	2.317932655
ERCC2_2	4.775107625	5.491218166	18.91710668	17.95868998	7.69470753	7.991636046	14.66536689	15.98633331	2.396436139	2.308928832
ERCC2_3	4.602053617	5.337311076	18.14297956	16.82632178	8.161644042	8.097381705	14.42623127	16.09281405	2.186444347	2.141861717
ERCC3_1	4.180079466	4.460571158	18.31011455	15.72812111	8.414714192	8.718666435	14.89556769	15.98308812	2.248310966	2.183778997
ERCC3_2	4.406682776	4.909585534	15.44012147	13.20125368	8.501245399	8.336394477	14.17595993	15.08707772	2.612115127	2.24244513
ERCC3_3	4.281245802	5.104961833	16.91945681	14.73123979	8.503835944	9.137296302	14.56229776	16.12117416	2.260347746	2.079686521
ERCC4_1_1	3.727739595	3.929850483	13.50337164	12.74756708	7.474661445	7.494658406	13.93446521	14.64166361	1.873731573	2.124399248
ERCC4_1_2	3.366293122	3.924600441	14.12946882	12.2075646	7.992537913	7.886712171	13.87986141	15.40342167	2.122844971	2.194441753
ERCC4_1_3	3.073162483	4.285307867	14.68779034	13.50307957	7.394754352	7.780943117	14.26714449	14.92745237	2.190005898	2.087967786
ERCC4_2_1	4.196708715	4.535576304	15.12201837	14.1311014	8.75838742	8.436678198	14.10685519	15.27322908	2.267482069	2.090157349
ERCC4_2_2	4.09475075	4.874505171	15.65398958	14.3494703	8.347757869	8.504520003	14.09593498	15.19386604	2.095466787	2.192654608
ERCC4_2_3	4.119877038	4.671648414	16.61802075	14.29569665	8.012773825	8.513639045	14.55830728	15.45502112	2.327097098	2.339176597
ERCC5_1	4.161454984	4.86786319	16.06117598	14.38605004	8.105574385	8.513937252	14.07721967	15.56137244	2.196911288	2.232498161
ERCC5_2	3.905953976	4.455800463	15.08191011	13.64622806	7.773643883	8.691786722	13.91222554	15.54807663	2.292045398	1.915418166
ERCC5_3	3.93571027	4.43842579	15.02474413	14.1592496	7.857180473	7.978020449	14.02858761	15.52430417	2.282639606	2.177058949
ERCC6_1	4.376522268	4.698806034	15.32952811	14.40058068	8.446291514	8.572361025	13.8862368	15.11810148	2.101767024	2.005412457
ERCC6_2	4.325005249	4.698195395	16.10273474	14.54417079	8.765605113	8.873620586	14.08924127	15.78856223	2.016606723	2.304897229
ERCC6_3	4.108926845	4.997807979	15.77591481	14.34772957	8.701336334	8.665767328	13.72657911	15.44228602	2.156666863	2.32236175
ERCC6L_1	5.364635787	6.054602927	18.53686185	16.71577769	9.490914772	10.17035226	15.20520114	17.50863414	2.272263735	2.183470976

ERCC6L_2	4.716827158	5.335863923	16.69824236	14.880208281	9.590353428	9.273851713	14.64108481	16.38620402	2.158029965	2.166195256
ERCC6L_3	4.316340163	4.878842845	15.81411258	14.06535417	8.65168441	8.851934101	14.91611124	16.08017702	2.49571806	2.302172996
ERCC8_1	4.439174192	4.952480096	16.35996001	14.74846732	9.026325694	8.419406459	14.15990132	15.6245983	2.359521666	1.956216631
ERCC8_2	4.504904533	5.015238327	16.7256616	14.569014227	8.525570212	7.749122916	13.58993653	15.51113606	2.238840313	2.169032445
ERCC8_3	4.503219582	5.017406804	15.16598662	13.67979971	8.78447872	8.609034064	13.89006119	15.40504233	1.98210749	2.063262267
ESCO1_1	3.874440925	5.256633467	17.26309602	13.89633529	8.615498095	8.807359101	14.60968041	15.63360629	2.247402677	2.016123133
ESCO1_2	4.44136609	4.956065883	16.30485655	14.44395749	8.631959575	8.552748886	13.8596161	15.22770202	2.39756897	2.081730313
ESCO1_3	4.562034661	4.813643814	16.70391985	14.37670978	9.064789976	8.902665129	14.3351	15.62654319	2.234493666	2.119430693
ESPL1_1	4.417899126	5.210544334	13.99688122	12.79608974	8.988461423	9.269346213	15.28192364	16.13677997	2.490573025	2.277994926
ESPL1_2	4.3064941 01	4.92607392	14.71005759	12.81598216	9.027247813	9.462621071	15.35998574	16.45127811	2.514132985	2.401195916
ESPL1_3	3.946932007	5.031751837	15.51875442	15.09921005	9.121682524	9.273497881	15.02352227	16.1125539	2.684741744	2.417616817
EWSR1_1	4.722566166	5.286005369	17.08477828	16.1806011	9.041881204	9.001169313	14.84847261	16.47369694	2.296127365	2.463190305
EWSR1_2	4.386524949	5.476298614	16.95692027	15.34777028	8.890024771	8.860523836	14.0115444	15.87899911	2.231931356	1.862759837
EWSR1_3	4.672816754	5.804573071	16.1537353	14.26449401	8.919505325	9.268056331	14.45629483	16.4353176	2.252614642	2.222790607
EXO1_1	3.96520146	10.94062326	26.86661287	23.74740453	12.23718711	12.47762721	18.95531509	20.40188322	2.586660245	2.469519484
EXO1_2	9.65151 89414	10.84135566	25.01334016	21.80175366	12.46681892	11.69325044	18.60369956	18.90784723	2.509441775	2.490537472
EXO1_3	7.053008155	7.958688454	23.45734979	20.39715256	8.120282592	9.087632285	15.614133 48	16.3774161	2.484493218	2.105101151
EZH1_1	4.030690161	4.790434605	16.26030759	14.99880276	8.030435888	7.946081417	13.5943745	15.00558661	2.183358816	2.073268921
EZH1_2	4.626912101	5.106845663	15.31654808	13.79179095	8.432039904	8.47726491 2	14.44597605	15.28034789	1.896274628	2.463679891
EZH1_3	4.541730237	5.220601617	16.3700732 7	14.81635277	8.769406564	8.694682749	14.18569457	15.21888931	2.120602326	2.098954516
EZH2_1	4.235014835	4.857113608	17.06642971	14.00531125	8.107378427	8.641340611	14.79214682	16.46513301	2.263246146	2.044633478
EZH2_2	4.537324045	5.168126883	17.30508574	14.69665845	9.312046468	9.454079435	15.85081231	16.19195046	2.413563632	2.246026261
EZH2_3	4.324636167	5.091650084	16.98947148	14.44949705	9.855666651	9.955870547	15.32556889	16.11055823	2.338151665	2.269197784
FAAP20_1	4.117367868	5.297317935	16.21749971	15.36096541	8.744187084	9.123026636	14.73044876	15.69755838	2.018761 58	1.904586931
FAAP20_2	4.402537526	5.363095666	16.22267341	14.40759561	8.818633477	8.967166647	13.75882815	15.86602507	2.109449788	1.907904174
FAAP20_3	4.264326567	4.767902278	16.36221424	14.3983782	8.500935345	8.605119616	14.68335367	16.14644026	2.365243054	1.871555105
FAAP24_1	3.654824749	4.691350949	13.66218069	12.47908909	8.650038807	9.176176333	13.50661614	14.81751756	2.152169921	1.952985089
FAAP24_2	3.930405231	4.721239669	14.37036756	13.45223601	8.30416516	8.661046676	14.16873612	14.94710587	2.26481906	1.861693509
FAAP24_3	3.920657713	4.644508619	14.3550575	13.25825763	8.057910461	8.557678633	14.08040593	15.16902126	2.195736002	2.04323072
FAN1_1	4.889191954	5.569322195	17.03327426	14.69616943	8.946573923	8.668124442	14.73050073	15.58401603	1.895750186	2.33642421
FAN1_2	4.88975107	5.367255436	15.86403628	13.92076909	8.759737467	9.352656389	14.11914668	15.60908053	2.349638571	2.269500498
FAN1_3	5.154960119	5.575182087	17.33658292	15.81398751	9.139938944	9.724335765	14.44098094	15.77242718	2.175449461	2.289717183
FANCA_1	3.619370383	4.144335192	13.81122571	12.03752365	8.104076863	7.941895194	13.59001509	14.98457053	2.219458375	2.164106662
FANCA_2	3.402232462	4.107993215	14.90448765	12.72144343	6.741886127	6.793153279	13.64712646	14.55252685	2.286367534	1.902417188
FANCA_3	3.638281325	4.805875193	13.99846253	12.85616681	8.374901459	8.689580136	13.97038472	15.50649171	2.187891651	2.182605153
FANCB_1	3.556051844	4.595501184	15.2414516 5	13.4633411	7.498689831	7.369709689	13.58857455	14.86262255	2.06057121	1.768923305
FANCB_2	3.532431019	4.020462568	14.19405594	12.3966318	8.266290187	8.446506047	13.537927 62	14.68501679	2.332974894	1.991590104
FANCB_3	4.038303229	4.45460223	13.63396699	12.24479503	8.173823375	8.409907848	13.73603811	15.41276776	2.225145459	2.119081106
FANCC_1	3.913226528	4.525731085	15.83586315	13.00710578	8.092946588	8.247066214	13.93322371	14.39989394	2.156606913	2.378766262
FANCC_2	4.042251734	4.762559568	15.19876878	13.56392687	8.273424286	8.604590732	14.32711886	15.11421607	2.143506996	2.032188906
FANCC_3	3.946222754	4.552956667	16.75904599	14.55401637	8.378911504	8.572754755	13.95448222	14.62023791	2.287246496	2.319635486
FANCD2_1	3.426187993	3.986382469	12.89053964	11.83766522	8.557602044	8.12189346	13.57030655	15.31411666	2.063930091	2.266302449
FANCD2_2	3.413955831	3.993226939	12.9657333 1	11.64333658	8.556566077	8.516998017	13.20339471	14.29704636	2.381036383	2.002585444
FANCD2_3	3.627889026	4.107206625	13.87096638	12.70364363	8.362313466	8.248807602	13.27974246	14.42540656	2.025267178	1.885129079
FANCE_1	4.143425536	4.345273046	14.76602103	12.61548647	8.363425965	8.510218035	13.81076534	15.22237894	2.187288608	2.170651664
FANCE_2	4.315668283	4.76270315	15.03142443	13.20601829	8.148139371	8.594839313	13.23313429	14.44079525	1.841729914	1.829327467

FANCE_3	4.212824158	4.348712604	14.02407117	12.77977054	7.993094033	8.534191496	13.72117823	15.04855672	2.184397163	2.168463274
FANCF_1	3.634947731	4.247969125	13.50242201	12.26738178	8.659941297	9.062333123	13.87317795	15.51263555	1.908053987	1.76268398
FANCF_2	3.549871259	3.846463036	15.8754769	13.908325	6.377209627	6.748381129	13.24416667	14.59275858	2.035087062	1.972771781
FANCF_3	3.856556177	4.637140834	13.23532869	11.47510296	7.944678231	8.375300314	13.39149522	14.17513863	2.456838677	1.796065343
FANCG_1	4.845774567	4.921553127	15.05115839	13.34718538	8.626103463	9.096985605	14.21659278	15.75468419	2.167033655	2.316490517
FANCG_2	4.24504464	4.537395552	14.63858592	12.81333848	8.733829263	9.074736575	13.83828186	15.03935191	2.227091004	1.823803279
FANCG_3	3.733756204	4.708660125	15.35793092	13.02225396	8.40884063	8.329330387	13.99411126	14.92513972	2.119755013	2.330765131
FANCI_1	3.523379019	4.187605958	13.72740787	12.17363662	8.056834381	8.262480479	13.75641299	14.85598793	1.976913682	2.063463061
FANCI_2	3.427677429	3.774281008	11.82334179	11.93175545	7.698067367	7.918664416	13.49736162	14.88367447	2.362094589	2.164122269
FANCI_3	3.630879412	4.052540783	14.1824208	12.33674583	8.309103974	8.315460587	14.09521093	15.40304597	2.017692078	2.22639737
FANCL_1	3.970892907	4.54463097	14.19067176	12.59091382	8.399150013	8.574597468	13.4676717	15.03828849	2.07912477	2.124123298
FANCL_2	4.01411464	4.3527405	15.61626428	13.32867033	7.446323536	7.779265414	13.40032474	14.43114321	2.21156143	1.911667765
FANCL_3	3.772747446	4.987719615	14.37314431	13.10504761	8.503842015	8.151900588	13.10032495	15.06191534	1.781352545	2.174294916
FANCM_1	3.43477886	4.28764732	13.76094742	13.11024988	8.372205672	8.624719744	13.45618228	15.49376202	2.262949348	2.357799777
FANCM_2	3.856203892	4.376609388	13.84219186	12.2928626	8.132429465	9.03664503	13.72136705	14.9680606	2.024385704	2.039593549
FANCM_3	4.422679256	4.96371793	15.11368925	13.94976819	8.221232023	8.606475716	14.25275748	15.54327124	2.469458734	2.047517269
FBH1_1	4.455185245	5.088244664	16.67862376	13.53174105	8.328656108	8.668669655	14.42916693	16.35435955	1.906152723	2.114088575
FBH1_2	4.359008004	4.85142177	15.72891909	13.78967029	8.290493298	8.414626648	13.96274234	15.26972624	2.129534325	2.03236154
FBH1_3	3.96319188	4.563168119	16.97315398	14.15870562	7.767193564	8.053450027	14.01473383	15.50237954	1.999406117	1.814152174
FBXW7_1_1	4.241160774	4.585589643	16.07771078	13.64357967	7.354498765	7.521301504	13.29216402	14.67514434	2.402013976	1.856453317
FBXW7_1_2	4.125888325	5.207100592	15.7316552	12.63306275	8.476299566	9.397813163	14.2190541	14.13780925	2	2.117158515
FBXW7_1_3	4.580786118	4.990535769	15.71770348	13.78170332	8.857752243	8.668147955	13.43466328	15.23403384	2.237900714	2.0526824
FBXW7_2_1	3.750275149	4.78533742	15.75133645	14.08357514	8.138309641	8.279121541	13.96757679	15.10757168	2.223391658	2.306725533
FBXW7_2_2	4.288771699	5.212683134	16.21937328	14.217293	8.017396237	9.027518788	13.7129167	15.76885243	2.337547055	2.153649553
FBXW7_2_3	4.359207142	4.602734075	16.19518697	13.54878255	8.078452388	8.243038565	14.34597132	16.05510001	2.086012287	1.727784681
FBXW7_3_1	4.085618665	4.54272373	15.14696113	13.94041527	8.280826106	8.273619582	13.54639537	15.67339103	2.342846833	2.391117774
FBXW7_3_2	3.955473769	4.940517029	15.70842109	13.41761618	8.233499016	8.497312816	13.77025636	15.37881585	2.155927993	2.320116006
FBXW7_3_3	4.580862353	5.027405547	16.67725239	14.28670558	8.571700886	8.935492631	14.03012484	15.62157437	2.258346443	2.240967601
FEN1_1	2.690668992	2.969669077	9.1695572	7.624868336	7.099012215	7.084049777	11.23442571	11.5999841	1.962377661	1.7656524
FEN1_2	2.80578393	3.252998458	9.617321249	8.945680247	7.353669966	7.633599984	11.95376374	12.85260152	1.84158033	1.829631838
FEN1_3	3.125772856	3.344303884	10.47261252	9.56143119	7.550803816	7.638616362	12.03568725	13.35599677	1.724557195	1.870141911
FIP1L1_1	4.725328412	5.67847528	17.06316029	14.71162937	10.07881326	9.959224924	15.6471459	17.49691661	2.879440258	2.757606737
FIP1L1_2	5.494386406	6.055502978	17.79981433	15.88562351	10.65663215	10.94473045	17.1034175	18.21145866	3.053949979	2.87400381
FIP1L1_3	5.922147056	5.177321332	16.05803446	12.79922429	9.553047237	9.269189102	15.88356821	16.30971567	2.90067715	2.678698662
GEN1_1	4.226792191	4.235567515	15.55191056	14.83150965	7.326899591	7.403313094	13.85272707	16.06744677	1.973873367	2.264107793
GEN1_2	4.552794957	5.395736848	17.46488248	15.27896207	8.207388816	8.857912488	13.8351634	15.18747778	2.301252515	2.078355609
GEN1_3	4.69292206	5.135556663	15.87688557	13.14939463	8.507385167	9.055546271	13.66458952	14.87317348	2.515780598	2.234500527
GPN1_1	6.262199089	6.137203449	20.34380539	17.67743125	10.25120348	10.19178178	14.68459152	16.03566824	3.444264304	3.437885688
GPN1_2	5.809181841	6.682191781	20.84165225	18.65434049	9.518693387	9.176086985	15.12788344	16.74780025	2.773662769	2.808953275
GPN1_3	5.79188084	6.057261004	19.40867457	18.00269039	9.834203456	9.933245787	16.52870016	17.37342603	3.401313409	3.172964502
GTF2H1_1	5.659480967	6.354191634	22.56284458	19.52331606	7.857334406	8.901668314	15.60233314	16.70056454	2.480919956	2.577941906
GTF2H1_2	4.261231991	4.744094962	18.54171033	15.91792543	8.209216261	8.17830635	15.22234666	15.6704925	2.16342558	2.333409233
GTF2H1_3	5.264788305	6.247314439	22.82816491	20.25905229	8.227549308	8.408025013	15.57670138	17.09506132	2.471111281	2.052918379
GTF2H2_1	4.587444985	4.67117394	15.93665656	14.24503077	8.509099711	9.109289171	13.94366868	15.42744157	2.119541476	2.2046965
GTF2H2_2	4.716443784	4.541849632	17.5212546	15.93300345	8.008552277	8.527896323	15.93028379	16.80709901	2.731074568	2.225444232
GTF2H2_3	4.008651619	4.44428066	15.74306857	14.48476131	7.412374748	7.472407472	15.16666906	15.70849019	2.384412787	2.272357061

GTF2H3_1	3.955176141	4.775445579	16.64533703	14.63899943	7.765497292	7.945230233	14.003594021	15.10333149	2.33625778	2.266201703
GTF2H3_2	4.321303996	4.761333784	17.6464361	15.37903915	7.355139819	8.330848945	14.03586451	15.77055564	2.218243943	2.238014053
GTF2H3_3	4.636493517	5.286181237	17.86662909	16.37619874	8.69253067	8.997792494	15.74959672	16.68726063	2.682952022	2.412502529
GTF2H4_1	4.802238966	5.369038672	19.46032528	17.94820717	8.168574402	8.448500424	14.96017569	16.06563117	2.073489101	2.190475217
GTF2H4_2	5.271209601	5.705195128	21.90195927	18.83565301	7.697128711	8.216656944	14.93184468	16.41999418	2.293635129	2.232215729
GTF2H4_3	6.355698691	6.596154848	24.39736156	21.13286115	8.398179784	8.927481344	15.53066443	17.95409253	2.264925302	2.436094657
GTF2H5_1	4.232033568	5.024388061	17.40442656	14.99384262	8.118370711	8.341210035	15.17796659	16.66037498	2.251761265	2.083995845
GTF2H5_2	4.290027011	4.703708967	18.63024196	14.65104432	7.354876631	7.833663965	14.90505761	16.36316233	2.275159364	2.350477996
GTF2H5_3	4.458397744	5.406365495	16.98212226	15.47954434	8.079987253	8.270305196	14.23585665	15.50678153	2.188612364	2.396894165
GTF3C4_1	4.827248223	5.589982999	19.93632425	15.29153135	9.850787295	10.02603104	15.6323788	16.92596711	2.765306122	2.803278888
GTF3C4_2	4.352350607	5.652230616	16.4885149	15.96576261	8.303153635	8.895411982	15.74616489	16.54705903	2.620114513	2.07089435
GTF3C4_3	5.110383414	5.756478202	17.32174656	15.77878117	8.641013677	9.126472974	15.5681618	17.33820063	2.556857957	2.692240014
H2AFX_1	4.543893457	4.747822443	16.61502011	14.66834241	8.314677509	8.830349281	13.73143323	15.49958101	2.304548766	2.302503347
H2AFX_2	4.510469029	5.153481614	16.41862796	15.26920289	8.461495555	8.645244631	14.40914645	15.89927146	2.132064736	2.303131278
H2AFX_3	4.954835836	5.084384248	18.02828181	14.90382302	8.332019476	8.855099361	14.07521835	15.72847569	2.285965491	2.224066827
H2AFZ_1	5.254303409	5.52201691	17.70353495	15.62071533	8.767935587	8.535629167	14.14455811	15.70550558	2.178225015	2.264087279
H2AFZ_2	4.790737348	5.665210185	18.05309735	16.7730020	8.601013029	8.427074013	14.35125345	15.69456461	2.108095599	1.959571564
H2AFZ_3	4.899971865	5.741941375	17.28675525	15.36754322	8.279162477	8.173649902	13.97518856	14.54004448	2.595631859	2.146467854
HAT1_1	3.967974718	5.955445105	15.24728476	14.17993542	9.227283831	9.438286432	14.19917553	16.01504689	2.25434113	2.009143095
HAT1_2	3.921915755	5.075252714	15.74066278	13.91942859	8.891728077	8.836833461	14.13043478	15.59630528	2.252082012	2.195032722
HAT1_3	4.179029673	4.660198711	15.39632575	13.62857421	7.961565138	8.622292553	13.7708178	14.78513113	2.174910065	2.090608187
HDAC1_1	4.17051299	5.239539705	15.77545287	13.40260562	8.365502626	8.762728247	14.57464747	15.71812546	2.443952394	2.111553785
HDAC1_2	4.359299257	4.499541537	14.78063353	13.12003108	8.50795909	8.431197049	13.92403524	15.30222783	2.207246744	2.286374573
HDAC1_3	4.192704616	4.961663573	15.95809782	13.85746308	8.317127111	8.533461546	14.6387076	15.8417757	2.204117401	2.516371634
HDAC2_1	6.034588495	5.868442512	20.06325474	19.03724981	7.468867817	8.415623197	14.69649405	14.65347402	2.051745192	1.865265836
HDAC2_2	4.581687875	4.722132991	16.90140845	14.73688655	9.196061224	8.645659375	14.55930544	15.68539542	2.044547162	2.123796245
HDAC2_3	4.921111946	5.361860152	17.70199547	14.61778848	9.048076729	9.074500032	15.49802492	16.6480576	2.603394693	2.284735371
HELQ_1	4.322059768	4.877806214	16.92601	14.34441744	8.341653245	8.24542118	14.30775954	15.28656553	1.603744966	1.676888378
HELQ_2	4.836281521	4.255803368	13.86744209	12.2748498	7.419691665	6.528908869	12.86463799	15.001254291	1.821251961	1.659410431
HELQ_3	4.78553568	5.677475884	16.3763861	15.50861795	8.544609817	9.375662048	14.59896281	16.10562205	1.584541581	1.57193811
HIRA_1_1	4.250185598	5.49789484	16.26364526	16.1628342	8.705980914	8.932910408	14.30910158	16.14063721	2.235198091	2.217273826
HIRA_1_2	5.006399344	5.521132663	18.07890183	15.7165415	8.560660445	8.626923789	14.39269666	15.08062897	2.20360485	2.115119233
HIRA_1_3	4.415927785	5.284910902	17.42140657	14.67725182	8.693391726	8.822290109	14.31841609	15.49509316	2.035234954	1.961632954
HIRA_2_1	3.98379662	4.297717646	14.50486594	12.76146965	8.084419329	8.69406332	14.3350751 5	14.77834168	2.21096021	1.78458521
HIRA_2_2	4.706151507	5.243410401	16.13956335	15.14123472	7.492393449	7.666011743	12.05223813	13.5009527	2.191214991	1.615878212
HIRA_2_3	4.366320176	5.103527252	16.21489422	15.04112083	8.617659562	8.801418513	14.65491031	15.59137371	2.136154161	2.114781415
HLTF_1	3.57307877	4.100506542	22.36622809	20.81615803	7.069272969	7.347913986	19.32889552	20.36229234	1.329444975	1.403295555
HLTF_2	3.535056662	4.470371296	22.11401993	20.20371691	7.282126201	7.287067974	17.54512769	18.73849411	1.547494477	1.287566913
HLTF_3	3.897268549	4.626258873	23.30830554	20.93173937	7.563347263	7.834198681	17.28866586	18.62167863	1.579943039	1.458053871
HMGB1_1_1	4.711256291	5.252348406	15.22534397	14.79213492	8.781026125	9.302623982	14.58014776	16.24148979	2.152338614	2.010885424
HMGB1_1_2	4.164548547	4.640329367	15.26430317	14.7662551 7	8.273024753	8.516007772	14.14864853	15.49832653	2.221401522	1.848838044
HMGB1_1_3	4.150111887	4.279514081	14.66994139	13.03077174	7.726173614	8.647625333	13.62192926	14.91151381	1.933185376	1.915886703
HMGB1_2_1	4.152605809	4.873372398	16.01310478	13.91501963	8.656194728	8.874623475	14.06057752	15.2883474	2.059234588	2.230079821
HMGB1_2_2	4.803673608	5.265637749	15.55506539	14.4923119	8.532640968	9.050857345	14.07227075	15.66422667	2.157686413	1.946164107
HMGB1_2_3	4.06603248	4.761274263	16.09059488	14.72605353	8.4076053	8.278483649	14.21012162	15.37524366	2.250672048	1.963692026
HNRNPA1_1	4.413947401	5.178668121	16.71275329	14.64476401	8.025984833	8.359914168	13.76237839	15.11079921	1.966523046	1.838911364

HNRNPA1_2	5.015444438	5.463811025	17.22149842	16.32247359	8.562305871	8.951237526	13.896733771	15.58086513	2.087036294	2.136334516
HNRNPA1_3	4.677572686	5.374665477	16.98930793	15.71694114	8.261576093	8.687133692	13.49208868	15.29617697	2.203075443	2.279211019
HNRNPUL1_1	4.726240302	5.110846024	16.53619694	14.46546276	8.600350031	9.256635055	14.37414147	15.89240765	2.297997347	1.710326353
HNRNPUL1_2	4.616650462	5.366789431	17.43563013	16.82705146	8.227085748	8.403951038	14.07038749	15.77608422	2.27528204	1.985773874
HNRNPUL1_3	3.997803294	5.026124908	15.58219322	13.46044317	7.887050709	7.777449254	14.02111775	15.07549791	2.314272597	2.130883075
HNRNPUL2_1	4.828055325	5.593857153	18.82430344	17.34154362	8.092547406	8.797555614	14.57371351	15.69373802	2.033758181	2.307249919
HNRNPUL2_2	4.477020923	5.282122048	16.09160284	14.56153714	8.516282363	8.921519092	14.11382166	15.35181374	2.358374726	1.966894772
HNRNPUL2_3	4.365851163	5.516749944	16.40289396	14.40459039	8.574822166	8.207017047	14.08426941	15.5606173	1.841782879	2.16415183
HP1BP3_1	4.383229534	5.00177887	15.54649206	14.55986005	8.314091268	8.807964513	14.33500477	15.91496556	2.408529362	2.129263037
HP1BP3_2	4.320586195	4.978524623	16.2994443	14.33862632	8.213786931	8.119193103	14.17132264	15.59121746	2.397727692	2.198739046
HP1BP3_3	4.182818856	4.808656147	15.68572318	14.26373926	8.37191787	8.755589615	13.88522208	15.50874731	2.348049248	2.172433125
HUS1_1	4.845205675	5.038609417	15.37577097	13.70935327	8.708829233	9.112017876	13.78544501	15.50360733	2.304990598	2.355674592
HUS1_2	4.400870609	5.312435687	16.09156869	14.49400731	8.014134234	8.366851099	13.79574518	15.11231151	2.260910709	2.317703397
HUS1_3	5.303145131	5.568987562	16.40685931	14.71635601	8.846455498	8.794599579	14.46999132	15.88149898	2.410968246	2.247236462
HUS1B_1	4.567232347	5.386066585	15.85429183	14.20087205	8.505259423	8.941038632	13.938742	15.71926862	2.392703863	2.032004584
HUS1B_2	4.564785411	4.9624688	15.15137682	14.72517653	8.403950597	8.523510033	14.25805269	15.82946279	2.583088594	2.227272727
HUS1B_3	4.116648801	4.887394099	15.13577719	14.59103895	8.623934534	8.491223905	13.72601185	15.13924867	2.283014821	2.175798661
HUWE1_1	4.406859914	5.245726195	17.96834242	16.03312704	8.578929475	8.711975604	14.26176396	15.40657113	2.293281434	1.996970114
HUWE1_2	4.666167685	5.115581074	17.49567207	15.12012791	8.471907163	8.656290416	13.89774154	15.41512736	2.296166009	2.055051376
HUWE1_3	4.906084727	5.379412046	18.60017111	16.49270054	8.811525728	8.774381722	14.60250547	15.33729129	2.227915424	2.236342733
ING3_1	4.649355586	5.574662081	16.26644496	15.37202996	7.952799425	8.061879925	14.22657196	15.41348777	2.056187569	2.232926771
ING3_2	4.083714099	4.617913811	16.59492698	14.81344006	7.503298873	7.52201145	14.09128856	15.45084372	2.225743553	2.203681023
ING3_3	4.357333505	5.126289686	16.47962794	14.66526913	7.065642119	7.431728933	12.90454704	13.88544031	2.170159504	2.263779526
INIP_1	5.312667658	5.924427086	16.04663166	15.14064016	9.194973878	9.665233815	14.43317793	16.1022047	2.233038624	2.346477573
INIP_2	5.047451078	4.743530308	17.1858338	13.78647967	8.345035663	8.605221494	14.01718993	15.29585011	2.247533041	2.053447899
INIP_3	4.643945893	5.042098132	15.28531492	13.8981721	8.672055236	9.064942974	14.22145747	16.63461773	2.513882564	2.211134492
INO80_1	6.314897642	6.176484196	20.95268145	18.44003513	7.658085695	7.544906035	12.01541898	13.04226829	1.857251731	1.69320435
INO80_2	6.135478195	5.827870972	19.42256219	17.74960599	7.556143412	7.450763268	12.60612629	13.39018163	1.817854332	1.691011035
INO80_3	5.054988147	5.736602374	17.52953571	14.20737892	8.50653143	8.396302888	14.12401113	15.65061946	2.160648436	2.343017807
INO80B_1	5.480872713	5.911649333	18.75220523	16.39195915	6.620004593	6.42922528	11.1769463	12.50573058	1.924341756	1.545853331
INO80B_2	5.081333766	5.822601741	18.0852666	15.73547531	7.376275265	7.656632663	12.7062719	13.53483668	1.873688863	1.963074435
INO80B_3	5.619253896	6.290791781	18.98482979	17.19722409	6.717177352	6.734786656	11.74756437	12.45471014	1.749670352	1.542659925
INO80C_1	5.038119047	5.125666464	17.7813851	15.93525507	7.314370151	7.537577049	12.35668931	13.20341581	1.869669415	1.632605121
INO80C_2	4.695010156	5.229656642	16.77775281	15.06188777	7.874882705	7.70035461	12.57452423	14.09915868	1.759634475	1.967288355
INO80C_3	5.054555211	5.707163226	17.5545036	16.46900662	7.485768313	7.618083009	12.32514046	13.82400571	1.942998831	1.889550195
INO80D_1	4.27866896	4.772201998	15.95089886	13.93385088	8.772706281	9.232699284	14.7603277	16.19511091	2.259644015	2.162945266
INO80D_2	4.073874409	5.014967942	15.62108103	13.45714772	8.580409068	8.9473041777	14.09489425	15.68974063	2.333627112	2.333029701
INO80D_3	3.960112942	4.846727857	14.90599932	13.19790924	8.208865984	8.459266738	14.54921423	15.84712338	2.090692812	2.195376931
INO80E_1	3.4831182	3.71359676	14.75224015	13.25476763	7.236640419	7.431219082	13.10037456	14.59233703	2.236899469	2.157833266
INO80E_2	4.386542393	5.474017367	16.97052947	15.06688196	9.10521093	8.3853421113	14.22268716	15.74611797	2.319301698	3.00090027
INO80E_3	4.396047846	5.159225779	16.5800162	15.1458424	7.857273817	7.476622803	12.25586812	13.80235018	1.868806905	1.839132061
INTS3_1	5.245632023	5.905617196	18.16340513	15.56534596	8.604267724	8.798532466	14.43677051	15.82162471	2.21750313	2.259100152
INTS3_2	5.863728702	7.187024353	19.26731706	17.88129151	9.466600028	9.936987676	15.35175596	16.65236316	2.515738382	2.21792919
INTS3_3	6.426089649	7.156619873	19.92290555	19.24588401	13.37224837	13.09650634	20.69900022	21.36454362	3.613473439	3.320082796
INTS6_1	4.271707786	4.580633649	16.16760829	14.74547264	7.356785014	7.762675297	13.9474569	15.09309249	2.099671821	2.377327077
INTS6_2	4.562475861	5.030887746	17.31042139	14.24922001	8.469002976	9.386145214	14.13696364	16.06845167	2.332732569	2.202224364

INTS6_3	4.740460778	5.519385376	16.71207452	13.84924233	8.177368536	7.970471693	14.58331535	15.76423003	2.431538141	2.371814451
IPMK_1	3.844503152	4.542316802	15.70166048	13.53321662	8.56665085	8.701273295	14.2768193	14.63823243	2.312886092	2.477187734
IPMK_2	4.04772748	4.805309834	16.3471375	14.35301407	8.617333467	8.578880991	14.11522772	15.70123029	2.393599722	2.153615539
IPMK_3	4.266331764	5.014034638	15.04872346	13.03190263	10.22692773	10.40086674	16.58220427	17.47207589	2.557803144	2.36329301
KAT2A_1	3.84026981	4.831567378	15.7686851	14.10974953	8.668321373	9.1646249344	15.15923667	15.95757407	2.283804639	2.176246537
KAT2A_2	3.699813616	4.379134139	14.83477011	12.66858096	8.28277404	8.3055411379	14.00900351	15.30357313	2.11489276	2.236557957
KAT2A_3	4.129078464	4.928312653	15.64844305	13.68585351	8.462886847	9.975289922	14.58293623	16.55163266	2.16282928	2.160320243
KAT2B_1	3.649581301	4.467324379	16.25254812	13.80664746	8.057246719	8.634466959	14.82687595	15.68269191	2.041066503	2.179142042
KAT2B_2	4.365680897	4.629358581	15.2552377 4	12.79294381	8.365271294	9.100609001	13.79664331	15.45262846	2.116427547	1.959951971
KAT2B_3	4.563919689	4.669408803	16.42350195	13.84948565	8.583817745	9.014434879	13.9193729	15.63381425	2.16295577	2.429265504
KAT5_1	4.100142768	4.647542679	15.41944609	14.63358458	8.366506326	8.38118656	14.25351126	15.43165178	2.316606986	2.223369028
KAT5_2	4.018427654	4.634445389	15.84138171	13.84218144	7.928634804	8.358622855	14.25699645	16.00601905	2.030023073	2.387605518
KAT5_3	4.847927741	5.302009115	18.77205926	15.95866838	8.593623631	8.722792173	14.58418566	15.98993977	2.21008973	2.025020192
KAT6A_1	4.135321871	4.983843963	16.24519314	14.77207206	8.023461762	8.593463907	14.49693839	15.54251802	2.146695405	1.945544213
KAT6A_2	4.579995241	5.115508845	15.62706426	13.90547264	8.597991858	8.878303362	13.76632336	15.10283676	1.981249061	2.235854725
KAT6A_3	4.446174327	4.844490775	15.38304416	13.79775485	8.499246494	8.85164277	13.55497318	14.89991269	2.204958593	2.09009619
KAT6B_1	4.026396896	5.295873717	18.86099221	14.36532176	8.462686713	8.794580688	13.71901426	15.40580406	2.260380114	2.13739606
KAT6B_2	4.078405576	4.741031108	16.25240224	13.99611099	8.441637798	8.540538788	14.05216313	15.20757131	2.132478869	2.106537763
KAT6B_3	4.438741894	4.92005898	15.89384053	14.35707904	7.97201206	8.563603551	13.70413908	15.42228272	2.036952849	1.908325867
KAT7_1	2.618240617	3.284004167	11.81241615	10.57646991	8.292463183	8.998956101	14.36940457	14.81759502	2.164036658	2.296921006
KAT7_2	3.200531372	3.777359086	13.74756391	11.20141053	8.573602306	8.720925905	13.80190059	14.86053263	2.083426702	2.270918773
KAT7_3	3.476271565	3.682812728	12.26587086	11.36704521	8.695273632	9.49301508	14.25673027	15.03617024	2.040644827	1.966041106
KAT8_1_1	9.722103827	10.36507781	23.80927603	21.7702184	7.196931226	8.383771414	12.68631236	14.86295032	2.35339693	1.559157039
KAT8_1_2	4.642364251	4.656544159	14.84038473	15.09441511	8.341121495	9.182526528	13.93701128	15.21940856	2.191511507	1.969523038
KAT8_1_3	8.825726584	9.168890383	23.57555411	21.35890767	7.392392766	8.322653417	11.79320552	14.38905956	1.53818531	2.169190467
KAT8_2_1	4.48468488	4.867666532	15.57138552	13.72249553	8.608325148	8.534475266	13.69227262	14.91241137	2.162884685	2.253567583
KAT8_2_2	4.156516599	4.820234221	15.65356758	14.01800593	8.158730616	8.609637011	14.07332242	14.9959773	2.258520517	2.172616734
KAT8_2_3	4.047284766	4.962668977	15.20995591	14.32883135	8.424679176	8.568293163	14.55315083	14.49262912	2.542499254	2.220160594
KDM4A_1	4.505251673	4.962361019	17.03634094	15.34046294	8.782746567	8.772426521	14.96928999	15.91050581	1.922975845	2.291319591
KDM4A_2	4.30527995	5.143065188	16.30634221	14.06398724	8.575897549	8.673357102	14.47634519	15.28968324	2.188924437	2.281745139
KDM4A_3	4.347282948	5.034867683	16.91308536	15.21440063	9.109219325	9.416888949	14.27927025	15.91946013	2.557607957	2.28263249
KEAP1_1	4.275491135	5.105928974	17.00442165	15.36262625	8.056279913	7.999424157	15.31328717	15.91490992	2.501025505	2.3022343
KEAP1_2	5.114028236	5.3736356	16.43476731	15.42636568	9.257895933	8.531253503	14.71396543	15.72786386	2.144321541	2.161747366
KEAP1_3	4.492162016	4.972574305	17.07067743	15.85858909	8.229331116	8.556164808	14.10931654	15.36495386	2.08029875	2.146161264
LIG1_1	2.585325478	2.343116201	10.50048664	9.741614867	6.155595835	6.5754351	10.35985866	11.5158702	1.942837291	1.707977505
LIG1_2	2.806537735	3.036506115	11.43592483	9.365595771	6.543774997	6.629891603	11.32585501	12.10600433	2.030178019	1.905756899
LIG1_3	2.578916372	3.339032485	11.3689604	10.46557623	7.702241872	7.934650524	12.4843571	13.6625979	2.133478767	2.020628347
LIG3_1	3.399338259	4.302106616	15.90283883	12.97666611	6.760823973	7.302997687	12.66835521	13.81151643	2.029221421	1.866710749
LIG3_2	3.377419973	4.117850768	15.19031245	14.03800373	7.436293642	7.373422907	12.3039988	14.49206764	1.652553012	1.841217429
LIG3_3	3.411419856	4.076295065	15.24862206	13.7219665	6.996855586	6.6955044433	11.72615876	13.91688147	2.141884088	2.029630802
LIG4_1_1	4.365966324	5.116446222	16.0776782	14.45672123	8.555193463	9.001623639	14.45779385	16.00124102	2.513216317	2.093695507
LIG4_1_2	4.57170407	4.901205706	16.33428575	14.1507103 5	8.772118415	8.809447513	14.36137778	15.9559359	2.503457424	2.382919655
LIG4_1_3	3.837521568	4.434399135	15.36018195	13.77848092	7.867343484	7.984890316	14.15039645	15.71883902	2.3324894	2.438807293
LIG4_2_1	4.066219281	4.952862105	16.66810084	13.99155568	8.445677415	8.603233307	13.67797914	15.97219191	2.155300092	1.97972543
LIG4_2_2	4.319434707	5.09560161	15.99143855	14.252386034	8.337487537	7.85457091 4	14.06659235	14.59000131	2.33647019	2.516381775
LIG4_2_3	3.933961938	4.688860395	16.15446425	14.24927809	8.816693787	8.550419226	14.20376418	15.14784401	2.292636047	2.356153736

MAD2L2_1	4.716126061	4.939733402	18.19541037	14.95283894	8.040019146	8.770981175	14.41724247	15.63482731	2.587135695	2.503355457
MAD2L2_2	3.436653194	4.322153822	15.07563064	13.84210658	8.171549137	8.711555871	13.62794467	14.92566598	3.272191663	3.573502387
MAD2L2_3	4.06949203	5.007495387	16.67467369	14.6129008	8.238597372	8.390599421	14.0759016	15.73155505	2.424840904	2.216172972
MAU2_1	3.970379126	5.269290245	16.2192024	14.48965169	8.453844917	8.515467234	13.42067258	14.77168516	2.170222436	1.966072389
MAU2_2	3.685500533	4.39767795	14.56338241	12.53655506	8.072941367	8.230434933	13.56106456	15.12136461	2.331181142	2.186120748
MAU2_3	3.906694111	4.676895741	15.96202671	14.12733307	8.045854106	8.491144427	14.08221066	14.59328095	2.035092193	2.134150077
MBD2_1	4.476716116	4.713761036	16.09367644	14.1050782	8.387377148	8.930666363	14.4718816	15.2264901	2.18019642	2.240801504
MBD2_2	4.124109691	4.756499625	17.27773744	14.39263624	8.177413966	8.365202043	13.13722557	14.84671799	2.295643929	1.879767526
MBD2_3	3.876778902	4.762094803	15.87145367	13.84009827	8.844121311	9.227492352	14.18026459	16.12148304	2.26740516	2.325620221
MBD3_1	5.27454693	5.130440329	16.67985083	14.43619176	8.347173563	8.330364142	14.42322707	15.74255804	2.166622305	1.983217736
MBD3_2	4.622875922	4.989241179	16.3867046	14.38596063	8.273209136	8.871236154	14.42042445	15.48403796	2.269341222	2.246386692
MBD3_3	4.250318013	5.079722824	15.35291179	13.37450935	8.080627443	8.213619914	13.80255618	14.71527856	2.177283219	1.98512795
MBD4_1	4.060821687	5.251751691	16.85869632	15.40593233	8.551613565	8.438946895	14.20698229	15.72170039	2.363864093	2.18534056
MBD4_2	4.342615092	4.898825749	16.3884976	14.19359185	8.618019783	8.656864365	14.41021774	15.57663347	2.381714759	2.221926753
MBD4_3	4.650597652	5.75630658	16.42883336	14.88920368	8.539633707	8.8560331 08	14.17427415	16.12997533	2.056603293	2.182057775
MCM10_1	4.092950678	4.56475438	16.81736545	14.70038468	7.502781299	8.008779156	13.32677155	14.35393061	2.334499262	2.319537586
MCM10_2	4.020002724	4.507739776	16.46904444	13.90442744	7.561614413	8.233268697	13.47942678	15.45433955	2.165884264	2.149428325
MCM10_3	5.370783876	5.879059351	19.63707064	15.2916412	8.241367726	7.799496807	10.36412358	12.74230896	2.128600061	2.000764623
MCM2_1	4.46166393	5.241266827	17.34932367	14.31064816	7.112079535	8.726913112	11.16885062	12.59794829	2.176045115	2.115513768
MCM2_2	5.16418595	5.557402677	17.90685576	15.9594765 3	7.012529224	8.042564176	10.96079334	13.02073426	1.936542066	1.810892082
MCM2_3	5.249711554	5.423424456	17.62017128	15.49809131	6.699942143	8.079482837	10.16926318	12.23150077	1.916463743	2.017055288
MCM6_1	4.626777436	5.258235919	18.83008316	14.73665436	7.351559866	8.270240325	11.70494562	12.72251403	1.949203877	1.705852066
MCM6_2	4.679902956	5.567512261	17.89874693	14.42825259	7.586742241	8.375058896	11.34471916	13.7975404	2.029753288	2.000478583
MCM6_3	4.237357584	4.918986914	16.3036446	14.4093742	7.859224504	8.678780798	12.27663944	14.18481381	2.14713603	2.027344362
MCM7_1	4.666640401	4.466392099	15.82140887	13.52621531	8.339614133	8.174834563	13.33434803	15.25144032	1.929984538	2.286649865
MCM7_2	4.671898379	4.865401902	15.89371753	14.19870441	8.25128999	8.522452622	14.10328872	15.58436723	2.256671545	2.168542779
MCM7_3	4.259334416	4.856354612	16.13075555	14.59962191	8.207742871	8.135504431	14.02250063	14.90650177	2.190010822	2.287976108
MCRS1_1	4.289891001	5.171718968	15.35799017	14.01441749	8.47582264	8.304308566	13.9021771	15.56043069	1.905566268	2.199115149
MCRS1_2	4.478997695	5.302516866	15.83211771	14.41140722	8.577610397	8.777881108	14.32281478	15.55417874	2.347186096	2.242659996
MCRS1_3	4.312581697	5.114883537	17.11743296	14.87584293	8.175336869	8.436644504	13.81648622	15.82096394	2.116051021	2.161861214
MDC1_1	4.149113019	4.568636811	16.1937894	14.46622462	7.969097703	7.978636765	14.39447324	15.08333937	2.229139962	2.323790455
MDC1_2	4.322883409	5.080942163	14.04252728	7.961736117	8.483402768	14.04318942	15.08959281	2.177987495	2.054334337	
MDC1_3	4.748651396	5.26491484	16.18682977	14.42337686	8.208865586	8.446186448	14.06293537	15.49072736	2.129923392	2.208664184
MDM2_1	4.072429582	4.965357968	15.87004966	13.21203203	9.083335469	9.452431936	14.51789922	16.7549403	2.302887105	2.226851967
MDM2_2	4.292016506	5.368219438	16.02221933	13.57727176	7.823244797	8.799056688	14.06233961	15.96381068	2.154506167	2.114059676
MDM2_3	4.397659275	4.803450847	15.8521187	14.44605665	7.780110058	8.207367085	13.91833201	15.49345537	2.228319103	2.163507158
MEAF6_1	3.038792566	3.600226014	12.08722531	7.686319349	8.194516441	13.61110499	15.22352941	2.226641833	2.128102816	
MEAF6_2	2.865035699	3.00454234	11.87042353	10.69631593	7.233888858	7.704034858	13.06018731	14.28463657	1.993988225	1.734707343
MEAF6_3	2.568824978	3.130390341	12.37183595	10.2965626	7.240036628	7.322336279	13.51289953	14.70062324	2.096588317	2.025267465
MGMT_1	3.966076746	4.999947953	16.32237391	14.46117728	8.621217593	9.364569188	14.16321205	15.94296183	2.269518842	2.283020612
MGMT_2	4.905245447	5.14257393	16.51365189	14.0508942	8.626593859	8.64071507	14.37729487	15.69574013	2.296954722	2.188125019
MGMT_3	4.439662916	5.101053263	15.88098443	14.22978859	8.531385549	8.487781414	14.27421096	15.20104695	2.50844553	2.24328215
MLH1_1	24.52788782	26.18998791	39.7360653	37.16442622	13.72250075	13.51651 97	18.97835437	20.83294699	3.061656431	2.837294333
MLH1_2	24.00141284	26.04582639	42.02837688	37.80204272	13.51362408	13.01513628	19.03502177	20.60443358	3.115965388	2.905403595
MLH1_3	20.84316578	22.94873785	35.50072075	35.59391552	11.29053121	11.72741522	18.54667397	19.71844169	3.189060859	3.147181566
MLH3_1	4.201531383	5.233236782	16.76337143	13.92527119	8.547770925	9.0891446	14.57914641	15.69102239	2.399778055	2.128613926

MLH3_2	4.26448858	5.312400886	16.51479428	14.71380552	8.627582544	9.429768041	14.92466159	15.92131453	1.931589537	1.933054072
MLH3_3	4.321424957	5.021607236	15.71710894	14.55889514	8.673638708	9.065936892	14.23266133	15.74147121	2.240763542	2.393443788
MMS19_1	4.151882065	4.802648833	16.4293628	14.37698771	8.759983428	9.062273857	14.2311874	15.73410798	2.250310848	1.91510389
MMS19_2	4.590821816	4.61369983	15.76047359	14.24315583	8.1299767 16	8.670157869	13.82523279	14.62811971	2.236738539	2.110205929
MMS19_3	3.809255399	4.739113302	16.24472574	14.88000694	8.280153293	8.673984912	14.01721122	16.1812789	2.39044194	2.16499571
MMS22L_1	4.484508749	5.122624079	17.39436646	13.66463858	8.314327162	9.316756508	13.53848301	15.40637353	2.279307746	2.294583352
MMS22L_2	4.22832981	5.004687586	15.48547923	13.76601625	8.144163689	8.561120799	13.33524075	14.79391412	2.17417351	2.343021222
MMS22L_3	4.240103037	4.845618419	15.23448822	12.83594795	7.600151631	8.169416145	13.27140784	13.68197359	2.186632079	2.24792773
MNAT1_1	7.349856792	8.275363525	25.11531284	22.22409721	9.529486236	9.973128418	16.30826274	17.13846306	2.440492781	2.464997574
MNAT1_2	6.215617126	8.904052829	21.80952236	19.34675054	8.193338896	8.314506228	14.38136703	15.73216907	2.201889702	2.056411283
MNAT1_3	6.396515239	7.114544161	22.25804549	20.52792762	9.911787723	10.27119088	15.78665237	17.72420992	2.252269147	2.855984363
MORF4L1_1	4.2148427	4.914473054	14.49755149	13.26812113	8.223195533	8.462127986	13.5920019	14.30256786	2.19478738	1.95846828
MORF4L1_2	4.491995233	5.402290334	16.01382411	14.74986584	8.671029783	9.122414883	14.15187464	15.58243937	2.079520164	2.142560158
MORF4L1_3	4.073055898	4.864187669	15.30464593	13.91065953	8.360080774	9.007285238	14.11217369	15.49006914	2.27217473	1.99496397
MPG_1	2.134781448	3.013683452	9.366803283	9.327628494	7.904392803	7.732862564	12.27336761	14.61676218	1.948607212	1.908325034
MPG_2	4.006394817	4.434166694	15.3485751	13.92415228	7.969470693	7.998511152	13.3371829	14.949137	2.304680447	1.983296189
MPG_3	3.919399558	4.566410153	15.84999894	14.06961939	7.961185038	8.131350741	14.0935588	16.27311014	2.303652623	2.13195956
MPLKIP_1	4.713631312	5.259980256	15.85710545	15.16731236	8.422511531	8.862368899	14.62051356	15.54864456	2.139652264	2.251714266
MPLKIP_2	4.4084854	5.056786379	15.78098852	14.15907841	8.809271566	9.049475434	14.30047237	15.44662197	1.952715957	2.259193391
MPLKIP_3	4.393669631	5.140073082	15.81542393	14.41542622	9.323610135	8.556852542	14.65508473	14.17691117	2.449527873	1.990770066
MRE11_1	5.062983456	5.247795685	17.11236315	15.22655384	8.634723242	9.006281054	14.09346435	15.22680677	2.064471488	2.108542423
MRE11_2	4.36393568	4.776167443	16.33782556	14.2878999	8.217571095	8.538630944	13.80556902	15.62491176	2.163867337	2.1634264
MRE11_3	4.612789449	5.225630565	17.43274965	14.63133692	8.138159661	8.563953357	14.09994666	15.17203668	1.893782633	1.876172606
MRGBP_1	4.082560591	4.614110175	16.76633798	13.78319051	7.107011 24	7.432849645	11.78596392	13.39928591	1.816966836	1.647610412
MRGBP_2	3.985614712	4.681480931	15.31230571	13.45284252	6.142303878	7.152420104	10.61462792	11.48166466	1.497887594	1.429356789
MRGBP_3	4.087306384	4.917644115	14.65106539	13.8638253	6.377289297	6.819567811	10.31449779	11.56231992	1.632504988	1.286830772
MSH2_1	25.84022085	28.51131311	40.23851233	36.92486052	26.48885368	26.8687492	27.9517387	29.63375849	4.3524466	4.024734507
MSH2_2	19.93153462	23.61066212	38.67568361	35.13799243	19.35861349	19.78760792	24.41752376	28.04660363	4.525132914	4.223749488
MSH2_3	24.74936654	27.61342048	41.87381398	37.91091005	25.42136755	25.44123024	28.80731032	30.51203228	4.220186675	4.159716293
MSH3_1	5.064599463	6.015987266	18.68152666	17.58058457	8.390723276	8.570878559	13.88374073	15.38593123	2.565862351	1.766862575
MSH3_2	4.895446316	5.793005459	17.6497762	16.91366641	8.591540749	8.984927891	14.5856703	16.29847232	2.02655537	2.048446492
MSH3_3	4.444983944	5.499507263	16.60816034	15.13276391	8.160663851	8.851394604	13.67334514	15.08402034	2.084273092	2.111280088
MSH4_1	4.162333776	4.957654572	16.1585196	14.45990023	9.060181576	9.380418548	14.90804867	15.59663229	2.361278902	2.490235901
MSH4_2	4.414351757	4.587751856	15.76401132	14.24575731	8.471113516	8.45505618	14.85726454	15.0706671	2.092282986	2.267696492
MSH4_3	4.477314465	5.227720997	15.89590786	14.79468167	8.506860601	8.950583009	14.60876544	15.47556147	2.314365054	2.083780881
MSH5_1	4.48446581	4.86787285	16.03286541	14.04886491	8.468884911	8.904202594	13.86822443	15.07805824	2.018716404	2.286737921
MSH5_2	4.656907683	4.659631888	15.91097583	14.04745433	8.254030809	9.337036735	13.79816039	15.67600662	2.008671257	2.178509956
MSH5_3	4.837715896	5.476221514	17.16560891	15.78798631	8.655266951	8.849478457	14.46805276	15.79805859	2.216168158	2.232518955
MSH6_1	25.63032326	28.27811307	39.88632059	35.65586765	24.17004006	24.16269571	25.96819932	27.09089857	3.856222839	3.89872769
MSH6_2	25.13979544	27.38404668	39.57572639	35.70642643	18.06139514	17.63926072	21.98923771	23.48731707	3.583893146	3.413557259
MSH6_3	14.20016775	15.73859319	28.66644792	27.32364244	10.02721759	10.45266066	15.98747878	17.20786767	2.149373955	2.600518597
MTA1_1	4.575163399	4.81089143	16.20774477	14.02964661	8.361210712	8.458339449	13.99569913	15.60121351	2.294588361	2.057724346
MTA1_2	4.599691641	5.037761354	16.33968225	14.29381262	8.302931186	8.508368031	13.77202051	15.1222941	2.341585958	2.276107072
MTA1_3	4.436535488	4.955403965	16.86618381	14.64983404	8.126620868	8.765054569	14.13643003	15.28757267	2.323547697	2.48866319
MTA2_1	4.072756621	4.85718962	15.64144273	14.11683541	8.039962458	8.209704969	14.31884283	14.99973777	2.040789317	2.317182661
MTA2_2	4.041040261	4.864212971	15.66571932	14.05294413	7.906446922	8.087456191	14.64572056	15.29488271	2.121561824	2.164086612

MTA2_3	4.337899821	4.810655998	16.19389387	14.46338749	7.928575661	8.594637421	13.77990431	15.13786309	2.186084288	2.108802705
MTA3_1	4.37457027	4.739681726	15.85556866	14.37702308	8.600929436	9.109388655	14.5597475	15.60207752	2.301586171	2.224347979
MTA3_2	4.4424276	5.211001251	17.57955037	15.36169525	8.180134509	8.760874032	14.48742189	15.50123918	2.059772065	2.024224965
MTA3_3	4.015157315	4.760890615	16.09796343	14.48305701	7.899358116	7.766087094	14.27625518	15.28257294	2.08707739	2.186383822
MUS81_1	4.190127971	4.949593871	15.56234763	14.47202621	8.899443994	8.594191586	13.99056376	15.32650621	2.238336075	2.161513764
MUS81_2	4.19340055	5.32186819	16.31109106	14.95299422	8.783770925	8.846194607	14.64692324	16.34375028	1.950360691	1.878918113
MUS81_3	4.571032265	5.066418778	17.32806044	15.5817472	8.493275696	9.164574283	13.76858379	15.64752021	2.378776428	2.021240151
MUTYH_1	4.421652946	4.66153741	16.71848106	14.57753492	8.906124457	9.275472754	14.86859972	16.65752937	2.626072041	2.372949288
MUTYH_2	4.704365396	5.27418804	16.41767444	14.82777699	9.246932885	9.509737319	15.49453603	16.46983198	2.656434475	2.490332968
MUTYH_3	4.277915405	4.84749066	15.60916231	15.07356405	8.233724444	8.52856362	14.43008975	16.39854285	2.323310949	2.187377012
NABP1_1	4.426034365	5.209186803	16.32095695	14.24362372	8.43707826	8.805058806	14.05536107	15.54723118	2.353709908	2.731685292
NABP1_2	4.57141074	5.28271971	16.48814625	15.05981641	9.217888519	8.857189266	14.40008494	15.83935395	2.042194603	2.101526384
NABP1_3	4.264159195	4.900535105	15.13132699	13.96322806	8.547939336	9.187101108	14.41803519	15.48706658	2.145729966	2.177784974
NABP2_1_1	4.091656741	4.576591611	15.44596758	12.63515365	7.7534177	7.899749816	13.3134966	14.76375472	2.261114464	2.32913668
NABP2_1_2	4.006073094	4.60730676	15.36873025	13.45357905	7.624233499	7.691223298	13.58871315	14.50624244	2.019703288	1.97111247
NABP2_1_3	3.956810575	4.92531995	15.59132652	13.73445551	7.849953883	8.04442869	13.64717666	15.07228097	2.184631852	2.022906835
NABP2_2_1	4.232654173	4.721813767	15.43456709	14.15629616	8.895851173	9.289465079	14.60669068	16.40307068	2.484975253	2.167510354
NABP2_2_2	4.501774393	4.900773316	15.16139633	13.89658216	8.832324329	9.089288408	14.37213787	16.50871479	2.158152788	2.088600461
NABP2_2_3	4.056554088	4.821126067	16.35347382	14.88622587	8.501145846	8.55072135	13.62566558	15.57590817	2.072657351	2.166402577
NAT10_1	4.913024503	5.318882205	16.27429844	13.07337772	8.795662973	9.892878272	15.14997117	16.58443597	2.171582919	2.296629579
NAT10_2	4.588159831	5.260550822	14.81814381	13.62623875	9.005634808	9.787431453	14.26171032	15.70489215	2.207009298	1.853008717
NAT10_3	4.852637507	5.203307004	15.36289113	13.81722432	8.860328114	9.460645928	15.53798832	16.65571981	2.24795285	2.303834759
NBN_1	4.930044531	4.969922314	16.45580143	14.70483267	8.537000688	9.031592583	13.61876504	15.55852461	2.161239032	2.027672491
NBN_2	4.273810065	4.902827743	16.43227708	14.54035519	8.318211312	8.328474118	13.53446685	14.77294533	1.82198379	2.083856639
NBN_3	4.627614647	5.168645294	17.29615524	15.0254138	8.373994506	8.661200591	13.72837577	15.3558757	2.320962607	2.187178674
NEIL1_1	4.339006531	4.978714929	16.07405771	14.95766111	8.431945315	8.539061129	13.45023274	15.46719405	2.213284647	2.171229969
NEIL1_2	4.554666424	4.929584328	15.80411765	14.62716101	8.324169617	8.879728476	14.30211667	15.71854154	2.221041737	1.939437044
NEIL1_3	4.216045699	5.219006705	15.20037346	14.21478928	8.171320661	8.621668723	13.64740611	15.42020505	2.081981844	2.123382625
NEIL2_1	4.808362369	4.967018776	16.8422546	15.10904642	7.964423105	8.746439992	13.95382759	15.47327109	2.138666311	1.849967358
NEIL2_2	4.19361463	4.865204409	16.46575715	14.1219524	8.143685855	8.270496398	13.81225666	15.35368268	2.27052761	2.291323486
NEIL2_3	4.318888133	4.896810375	15.5825918	13.56362825	8.346937701	7.949130785	13.60565666	15.3004 93	2.313254537	2.177766209
NEIL3_1	4.281729426	4.787906097	15.40805254	14.13273088	8.348355961	8.611278237	13.90268799	14.81022291	2.207027796	2.489900559
NEIL3_2	3.950550927	4.825984233	16.25986113	14.06644347	8.366228125	8.873295164	13.48946864	15.27794733	2.292195221	2.159602881
NEIL3_3	4.72558945	5.161707347	17.26386223	15.01076191	8.812276824	9.437156704	14.33086108	15.5380092	2.375947513	2.206902596
NFRKB_1	4.426869351	5.003475138	16.10365796	14.59345513	8.648855081	8.826676317	14.13088725	15.41011242	2.357455754	2.312337267
NFRKB_2	4.870532203	5.544994494	16.8077 47	15.63821065	8.392575811	7.976093319	13.55399737	14.99280476	1.924195803	2.069312084
NFRKB_3	5.059810169	5.850613724	17.86313136	16.02721176	7.284257077	8.844777246	11.90916584	12.78650516	1.634422949	1.518119072
NHEJ1_1	4.408341077	5.193452475	16.45642555	15.29657399	8.690655121	8.332641508	13.94915356	15.33432468	2.21510708	2.265073604
NHEJ1_2	4.344282849	4.924707747	16.1867 3227	13.70130928	8.703375687	8.417545135	14.25608128	15.23008601	2.241777644	2.223651824
NHEJ1_3	4.462815291	5.272449711	16.35571768	14.94952623	8.393818603	8.927264745	14.47745034	16.09653214	2.233115468	2.22988276
NIPBL_1	4.067608095	5.179233925	16.00647601	13.80843461	8.500465014	8.548405345	14.42065426	16.19368421	2.498115341	2.156885668
NIPBL_2	4.15138375	5.009005693	16.56496083	14.30200704	9.119680241	9.183135705	14.36978207	15.82526476	2.155495087	2.430081377
NIPBL_3	4.004189324	4.808891404	15.10587426	13.71949735	8.876552745	9.341152393	14.40187497	15.45241741	2.417865707	2.020075524
NSD2_1	4.396993965	5.021644443	16.1021542 8	14.29710461	7.525400208	7.660719657	13.07658295	14.06861236	1.847911809	1.930411971
NSD2_2	3.903519124	5.076718947	16.3783387	14.37396964	7.847650163	8.263793145	13.64084349	14.11110772	2.157289574	1.802881146
NSD2_3	4.163104932	5.005436561	15.64095571	13.73223516	8.16155631	8.623392544	13.42633202	15.08821136	1.995971273	1.773487952

NSD3_1	4.380581495	4.977752108	16.162420881	14.253472121	8.552245748	8.641395612	14.336166434	16.07518745	2.185651839	2.119940722
NSD3_2	4.333982858	5.372345531	16.82410525	15.37001898	8.553612048	8.864508752	14.16789885	15.54164598	2.389062717	2.191790747
NSD3_3	4.256056863	4.841072416	16.68559657	14.47491498	8.400189322	8.916650508	14.03900648	15.64580635	2.267741134	2.322305929
NTHL1_1	3.906196773	4.402515723	16.44041764	14.64012829	6.821747307	6.977507998	13.37436564	15.36198611	1.983669603	1.806769885
NTHL1_2	4.457078165	5.263241332	18.22084669	15.88928227	8.45237667	8.410960767	14.46005392	15.49109904	2.23203255	2.257542541
NTHL1_3	4.11563425	5.064179388	16.12905556	14.37764533	8.627123637	8.848583461	13.68633131	15.8133749	1.880147579	2.179092838
NUDT1_1	3.200147557	3.706448165	14.67611576	12.25344536	7.119832687	7.205548737	12.97305819	14.17511888	1.877702962	1.896522691
NUDT1_2	2.862940637	3.567991591	13.3083551	11.26696696	6.67099941	6.736566253	12.39179174	13.60923278	1.963099416	2.173447078
NUDT1_3	4.15734321	5.012004395	16.31274373	14.22278991	8.173250693	8.289618256	13.93797751	15.30981499	2.451206486	2.140279904
OGA_1	4.249280216	5.330455418	16.20549222	14.32967459	6.864755349	6.144114603	11.36112313	13.05471869	1.880091009	1.473406658
OGA_2	4.837086727	5.400966334	17.38100543	15.19441894	8.517083637	8.81045378	14.11087021	15.74573523	2.396632865	2.241512346
OGA_3	4.823667502	5.314810168	16.26337631	15.41599426	7.741075967	8.457415951	13.60563754	14.93271178	2.146850934	2.103798297
OGG1_1	3.691657356	4.45297869	15.56801128	14.09351822	7.454377336	7.990509184	13.94378361	15.29606017	2.176537009	2.265164043
OGG1_2	4.738808416	4.799826901	16.46431903	14.70686394	8.115376979	8.290679417	13.38603481	14.47935367	2.097695805	2.010647646
OGG1_3	4.736547355	4.809133695	16.62946711	15.59816165	8.714842599	8.445886805	14.07386477	15.57109265	2.482051695	2.276457356
PALB2_1	5.27222825	5.459759273	17.94699314	15.88496438	9.585714585	9.057393384	14.88441642	16.51452262	2.595986709	2.62409514
PALB2_2	4.847978861	5.330167771	16.63198533	15.08114997	8.766268598	8.834763008	14.56316688	15.87275369	2.645440462	2.409617456
PALB2_3	5.042675133	5.584823464	17.21622015	14.9349576	8.927908662	9.090494782	15.2113070	16.53631612	2.265314215	2.362686689
PARG_1	4.756865022	5.267307996	15.35489317	13.90817673	8.365155412	8.883813115	13.80408426	15.86835715	2.213896768	2.002768263
PARG_2	4.181902091	4.822204466	15.47040118	14.18459796	8.556172863	8.689415506	13.89562016	14.62633152	2.077720977	1.79584613
PARG_3	4.614151591	5.32377335	15.92172484	15.08273282	8.493419626	8.931804022	14.22317072	16.02774423	2.130509623	2.194093672
PARP1_1	4.770670783	6.009161003	18.56134534	15.34181674	8.545102579	9.055497961	14.4092621	15.09978253	1.980649632	2.141772222
PARP1_2	4.871264744	5.782637897	17.97414123	16.56220506	8.622870745	8.648145551	13.67935476	15.20706311	2.382922636	1.94854589
PARP1_3	4.756407296	5.734857721	16.9579837	15.38303229	7.911509371	8.040286609	13.63083869	14.67956219	2.595361446	2.2792697
PARP2_1	4.087062411	4.55124368	15.39854884	13.32720958	7.810154394	8.079163124	13.47697352	15.13092557	2.13555724	2.016869033
PARP2_2	4.149697246	4.68593396	16.25147276	14.16741259	8.381947461	8.742995566	13.50649847	14.9585602	2.193810771	1.905623611
PARP2_3	4.184068637	4.57188863	15.10991998	13.35255731	7.926255928	8.145823323	13.20246869	15.56522488	2.063750772	2.162582836
PARP3_1	4.00059356	4.697433267	16.17277564	14.51580676	7.796200361	7.57073481	13.95838162	15.66872368	2.124694264	2.29095582
PARP3_2	4.070133165	4.66418122	15.7730105	14.0658441	8.630197436	8.813621708	13.82341678	15.38956167	2.289840352	2.170682748
PARP3_3	3.882891125	4.765928093	15.49391452	13.39741014	8.275892767	8.83367943	14.12587413	14.6048312	2.077882928	1.925317931
PARPBP_1	4.71593013	4.837781424	15.65828826	13.70960416	8.450169887	8.619093658	14.21896733	15.84782252	2.183539366	2.096076292
PARPBP_2	4.810473436	5.094574271	15.49057501	14.38992133	9.047157041	8.833697744	14.70925615	15.48421426	2.097741536	2.003477652
PARPBP_3	4.040228673	5.166065901	15.82113872	14.56317076	8.459895677	9.209605072	14.17281813	15.66872216	2.322584446	2.150994046
PAXIP1_1	4.662582899	4.925831148	15.72693816	13.84642281	8.91149006	9.691459105	14.50650241	16.43863754	2.452724209	2.319711738
PAXIP1_2	4.137818419	4.852108287	14.60036063	13.97252751	9.410367254	9.865727701	15.06101703	16.76127236	2.425472327	2.867728093
PAXIP1_3	4.408182432	4.950057456	15.98128587	14.94512677	8.391058855	8.99042462	13.67805512	16.24577771	2.217370719	2.199245973
PBRM1_1	4.516080326	4.957939705	16.04670057	14.56395172	7.691959193	8.017431376	14.28417967	15.15202232	2.067267046	1.924920174
PBRM1_2	3.590094023	4.36181141 5	14.86029482	12.89515814	7.186458546	7.319461192	12.63136927	14.35830651	2.048355066	1.585543087
PBRM1_3	4.367361854	4.883715151	15.06249309	13.30202789	7.274983081	7.24828144	13.038141	14.32764023	1.895147498	2.296883836
PCGF1_1	4.087285884	5.002484679	15.14646013	14.07638775	9.412572999	9.595968574	15.26859061	17.16750171	2.267656485	2.80394646
PCGF1_2	4.062502512	4.843218605	15.77961675	14.59532666	9.675195842	10.08176725	16.2133658	17.02146265	2.365224009	2.455491465
PCGF1_3	4.044049378	5.026490066	16.30784522	14.3133121	9.595182658	9.030027736	15.44608451	15.99418941	2.416778798	2.244480669
PCGF2_1_1	3.99090359	4.870360042	15.50642523	13.78741506	8.390479524	8.660257404	14.23022494	15.81327726	2.373426848	2.219176052
PCGF2_1_2	4.153424473	4.783036465	16.37548869	14.08323424	8.008953553	8.4830735	13.91463542	15.87582751	1.760162602	2.116550553
PCGF2_1_3	4.230132225	5.022710133	16.21279351	13.48563202	8.267596008	8.297264251	13.74367718	14.89042302	2.212381747	2.321486326
PCGF2_2_1	4.150083302	5.186234286	16.03917891	14.29472526	8.617934535	8.886531321	14.29499239	15.96235899	2.222080824	2.093007826

PCGF2_2_2	4.707765066	5.3357657631	16.53468201	13.75660664	8.28471323	8.543095584	14.24368542	15.9859251	2.045516863	2.352678933
PCGF2_2_3	4.712258281	4.843685516	16.45151554	15.38342165	8.890465281	9.099724335	14.20306406	16.09541041	2.439754992	2.249226478
PCGF3_1	4.892676768	5.1695119021	18.66148955	15.71298302	8.089208247	8.582933312	13.40156416	15.04465754	2.107296644	2.41537956
PCGF3_2	4.498441955	5.1692074931	16.83503905	14.37902796	8.682582903	8.849453057	14.56225132	16.2230756	2.144911572	2.27797445
PCGF3_3	4.279146039	4.77666888	16.24254708	14.4222437	8.03109635	8.365245224	13.9833662	15.33352301	2.216561626	1.927095655
PCGF5_1_2	4.181125327	5.1462295131	16.27124214	14.60650279	8.605045777	8.738365191	14.00052336	16.11666162	1.985525467	1.902531684
PCGF5_1_3	4.446220091	5.142512271	15.40129761	14.4009169	8.765555002	8.786459044	14.34936363	15.24646164	2.233655904	1.853301566
PCGF5_1_1_PCGF5_2_2	4.275038702	5.1085011113	16.94644762	15.2628639	8.150499424	8.15773668	14.16672777	15.2057252	2.309451542	2.307973849
PCGF5_2_1	4.271770566	4.8525521446	15.0807946	14.2717911	8.265810326	8.568512745	13.57429566	14.91185806	2.190853659	2.049735672
PCGF5_2_3	4.371609661	5.1338472421	16.58894406	15.44360695	8.300428431	8.393963408	13.80875729	15.39465611	2.318659875	2.038907714
PCGF6_1	4.114429862	4.4266230151	15.2499045	13.81362811	8.102745751	8.620664173	13.63602172	14.72245885	1.937549582	2.035584111
PCGF6_2	4.500650921	4.722767716	17.32809034	13.53790614	9.606730581	9.690217429	14.08444133	15.7660089	1.849315068	2.29786605
PCGF6_3	4.156059843	4.8416811251	15.65725131	14.1717611	8.096809519	8.466625121	13.61536246	15.61322257	2.265121617	2.300944168
PCNA_1	4.200332818	4.884592023	15.8723928	14.68557507	7.806503758	8.127508568	14.39269573	15.75452521	2.299635336	2.20897465
PCNA_2	5.156171854	5.7349810299	16.723197	15.1647161	8.490085698	8.73488634	14.35093628	16.01855197	2.210731603	2.223313955
PCNA_3	4.766495109	5.2805176898	15.86321296	13.47123271	8.730165392	9.211282641	13.88860572	15.6036474	2.192896744	2.172182689
PDS5B_1	4.454549548	5.442122214	17.39187393	16.30553815	9.126147371	9.24012868	14.39576581	16.03007599	2.211256421	2.194766126
PDS5B_2	4.577385364	5.5835430261	17.64905603	15.83206416	8.878409019	8.997729082	14.73686628	15.92075376	2.51692174	2.181471376
PDS5B_3	4.568438517	5.1198722171	16.91919236	15.51068864	9.249825567	9.513615913	14.68472402	15.93161474	2.293241857	2.266909577
PER1_1_1	4.035449686	4.9115572051	15.67225825	13.6945823	8.39468662	8.430868433	13.96985767	15.31900949	2.248511305	2.474683601
PER1_1_2	4.102822343	4.5129541491	15.99000191	13.9550120	8.647801585	8.550389008	14.35427675	14.91306914	2.069229747	2.397015498
PER1_1_3	5.214343488	5.6488352381	17.29711898	15.97894873	8.372480818	8.631431757	14.25113631	16.17856535	2.259307621	2.322547771
PER1_2_1	4.31944636	4.8545564944	15.74367089	14.4475832	8.174988619	8.281443481	14.18225873	15.34233956	2.320264673	2.358554307
PER1_2_2	4.614288978	4.965594435	16.20580834	14.165809	9.264524216	9.615684623	14.37944935	15.86126474	2.492799112	2.289284696
PER1_2_3	4.105791356	4.918173251	15.46156862	14.6643758	8.272572713	8.527833452	13.86558169	15.51838139	2.148285066	2.136578742
PHC1_1	4.229126214	5.1125162558	16.64309318	14.2776816	8.346437539	8.579507852	14.20996144	16.09554453	2.254692308	1.980568012
PHC1_2	3.962940123	4.594163247	14.88522616	13.7183440	8.433473473	8.281272603	14.28263765	15.66919508	2.341381508	2.190165315
PHC1_3	4.610919379	5.171356159	15.48917625	14.0042733	8.592327364	8.963590703	14.77436311	15.60739497	2.384986517	2.065913673
PHC2_1	4.177059462	4.920861269	15.26779747	14.2900627	8.226075137	8.252755523	14.1935562	15.70459803	2.17939103	2.175741861
PHC2_2	4.505227357	4.759445212	15.88510123	14.1865549	8.689983254	8.822194719	14.26869675	15.52612584	2.106472616	2.353096781
PHC2_3	4.18894234	4.9655508011	16.57980564	14.709765	8.635085656	9.243857533	13.98391755	15.3832041	2.183225837	2.381945797
PHC3_1	4.351410068	4.712825496	16.7374941	14.5409010	8.480242369	9.133637433	14.3877335	15.46259869	2.079375052	2.164145216
PHC3_2	4.277333597	4.828528185	15.38971925	13.5344586	8.185487816	8.508125085	13.73281062	14.96131841	1.991140053	2.385687965
PHC3_3	4.308100406	4.770050562	15.87932618	14.867522	8.19281026	8.55406277	14.19536777	14.82390569	2.248817627	2.483538134
PHF10_1	4.298257109	4.930630556	16.27966301	13.9727317	8.614291216	8.720392926	14.43024964	15.49268206	2.215504554	2.007462655
PHF10_2	3.536527028	4.1434547711	12.53837628	11.335655227	10.66704711	11.01732517	16.73556931	17.90147752	3.140604988	2.974353911
PHF10_3	4.190045501	5.0306696421	15.8558533	14.0995632	8.16178736	8.177067147	14.00830344	15.67812771	2.410671422	2.167993433
PIAS1_1	4.758473282	5.2548991761	17.49535844	14.798944681	8.960021101	8.864990044	15.02527231	16.10585123	2.248125969	2.314921177
PIAS1_2	4.460320817	5.430702846	16.27512871	14.5641542	8.596761076	8.662535602	14.24776925	15.76250065	2.423570672	2.120888041
PIAS1_3	4.615413633	5.2196863028	14.91592305	13.5996288	8.615693077	8.826476089	14.8549072	16.11642466	2.5496446	2.399298439
PIAS4_1	4.564474039	5.6793756568	16.6413813	16.26120535	8.645043237	9.011883491	14.71395703	16.26123696	2.480343176	2.157249967
PIAS4_2	4.806092646	5.327395637	17.2303774	15.4373973	8.507017458	8.592351317	14.60833985	15.53970552	2.074730842	2.259702196
PIAS4_3	4.155057019	4.993870867	17.84787813	14.4907217	8.040019146	8.604054773	14.26866616	15.08458146	2.281545753	2.056286562
PIF1_1_1	4.290207877	4.852728151	15.91507067	13.6446823	8.502544017	8.432931675	14.43140087	15.70436949	2.30771201	2.112559601
PIF1_1_2	4.736474346	5.7433654477	17.4774748	14.752367	8.799527101	8.899222958	15.01601852	16.14707222	2.471675717	2.422179778
PIF1_1_3	4.221640904	4.895118919	15.96118571	15.3997615	7.851576478	8.033345016	14.41585455	15.4045619	2.182059013	2.395443632

PIF1_2_1	3.600201735	4.39975506481	15.59380127	14.02670457	7.545589306	7.502429141	13.32726703	15.19551649	2.271661665	2.016164074
PIF1_2_2	4.097329209	4.424871958	14.96944947	12.58950498	8.289881713	8.514669122	14.02602334	15.56659582	2.237726988	2.086849942
PIF1_2_3	4.287012952	4.763997931	15.38188596	13.64108441	8.204554516	8.632002232	13.61467555	15.49745433	2.10399516	2.299629416
PLK1_1	4.996605567	4.802212538	16.34192332	12.79051001	9.00450795	8.102392387	14.93414642	16.2387841	2.472267316	1.994544835
PLK1_2	4.754442349	4.857513495	15.29744449	14.27294698	8.441288174	8.936644578	14.1713377	15.75702124	2.130748095	2.154617037
PLK1_3	4.116699461	5.538938739	14.35232677	11.63749749	8.529122913	9.578618806	13.96235294	16.888	2.348079728	1.505470466
PMS1_1	3.816063156	4.125536481	13.37984437	12.91763243	8.601347539	9.095723691	14.10078092	14.96843945	2.218410103	1.820154291
PMS1_2	3.855835994	4.372107349	14.24146986	11.82218096	8.633393863	8.714194988	13.76535049	15.54736378	2.053361897	2.045284958
PMS1_3	3.697061664	4.003328843	12.42716556	11.27930616	9.030273036	9.05658009	14.15238653	15.28531302	1.991276313	2.017621358
PMS2_1	16.74093506	18.37252858	37.80261227	33.06421601	14.15404872	14.26820867	21.14032264	21.89073463	3.679443873	3.211430672
PMS2_2	23.7962574	26.18191737	43.02760208	37.956197	25.01179979	25.18437248	29.65817491	30.64768073	4.391001989	4.142711551
PMS2_3	4.750447628	5.256746529	17.07836827	15.67015341	8.439149017	7.903537332	13.86136897	15.80253668	2.2496252	2.115009993
PNKP_1	4.628637646	5.015629577	13.65108921	8.790108493	9.0054792	14.09712428	15.87686623	2.053484763	1.958607558	
PNKP_2	4.200478749	4.750509616	16.47590452	14.24384052	8.676149096	8.981350664	14.42703288	15.64205247	2.170280171	2.051042903
PNKP_3	4.341388237	4.737140073	15.26865847	13.67229551	8.861641354	9.103213413	13.98777583	15.3556611	2.136967896	1.99738017
POLA1_1	4.067740329	4.94722219	13.96593674	13.09042561	8.501141263	8.974710125	14.2853646	15.56892305	2.407121345	2.624520065
POLA1_2	3.789055757	4.724323736	14.13441955	11.85510996	8.396000658	8.803547455	13.89689709	14.36134954	2.511136081	2.117438552
POLA1_3	3.689304661	4.283082591	14.90354952	12.785866	8.057945167	8.085663296	13.57993125	15.18991989	2.281219444	2.007123346
POLA2_1	3.743382908	4.425696676	15.01295534	13.543542	7.063908443	7.076860616	11.31109118	13.36371317	2.909008907	2.768350625
POLA2_2	3.975321124	4.51960866	15.46609925	13.53148677	7.800402775	8.074374924	13.29600305	15.04835369	2.228215133	1.951473653
POLA2_3	3.907822667	4.331529836	14.28726426	12.06741951	7.729910164	8.199527354	13.66925483	14.25395343	2.521022032	2.674397813
POLB_1	3.047897667	3.66769454	15.12626452	13.78090074	7.243867113	7.541339977	12.83459998	14.4461569	1.785367583	2.179621491
POLB_2	4.905828455	4.402628079	15.90713236	13.35046842	8.470663164	8.882831111	14.52631202	15.37482258	2.11306683	2.035132819
POLB_3	3.986245167	4.721770373	15.08120818	13.61692131	7.776013518	8.060477315	13.24957837	14.56881144	2.389308377	2.26168853
POLD1_1	3.418175533	3.795409362	12.97863765	11.39333597	7.670697752	8.477744144	11.94586324	14.07318317	2.215562285	2.006839453
POLD1_2	3.152212333	4.137873565	13.29937634	11.23813177	8.46462446	9.268622175	13.81967634	14.5182063	2.133589202	2.034356046
POLD1_3	3.528310938	4.180098861	14.35063524	10.62143927	8.072390125	8.674096448	13.35059345	14.2514955	2.21103605	1.856176421
POLD2_1	3.451024321	4.191846523	13.52428456	12.21315961	8.303675186	8.376981089	13.07662631	13.79746116	2.060762661	1.766580512
POLD2_2	3.231292517	4.297419457	14.60195628	12.74998666	6.548966012	6.747643734	11.86881632	12.73656532	1.452225841	1.92450587
POLD2_3	3.320075666	3.864174906	14.40264766	12.01537921	8.367402956	8.553308193	13.35977321	14.20766966	2.358675536	2.257138474
POLD3_1	3.334824224	4.151729484	12.94090828	11.37262123	8.532344213	8.834907637	14.42283745	15.40563967	2.130780233	1.932351821
POLD3_2	3.420026667	4.423684572	13.3531467	11.04242026	8.028884017	8.626139678	12.76356417	13.94740528	2.331546708	2.039028953
POLD3_3	3.825395371	4.726706866	14.23946349	10.61449135	7.978334721	7.968207153	12.07783916	14.32913258	2.00606785	1.91612865
POLD4_1	3.827250172	4.43374126	14.51580657	13.67710571	8.407376229	8.395959722	13.6023636	14.46158768	2.116844524	2.324290767
POLD4_2	4.513867817	4.887831724	15.42008416	13.74633962	8.206059672	8.832693043	13.93548193	14.94028178	2.158839403	2.016573964
POLD4_3	4.07581345	4.935782867	15.48049424	14.48970566	8.35683902	8.358357085	14.0155572	15.56803797	2.260034243	2.192225718
POLDIP3_1	4.627704869	5.396701817	17.95575631	14.45018326	9.202271016	8.499742224	15.12346503	15.61578689	2.123851569	1.997297046
POLDIP3_2	4.862399505	5.549195446	18.87406324	16.02089448	8.647600655	8.493555551	14.5880255	15.54801011	2.346691423	2.287426387
POLDIP3_3	4.185616228	5.492534649	15.07176169	17.15805608	8.800482218	10.28853684	14.94618624	16.20803716	2.214554032	2.05298462
POLE_1	4.576103461	4.08273658	16.47147316	12.99566227	7.601209742	7.802667357	12.58378333	14.75454299	2.490247206	2.095761134
POLE_2	4.266664544	4.905222303	15.02670517	13.42574393	8.13038690	8.236401164	14.18690142	15.54240352	2.187654252	1.842129333
POLE_3	4.517611026	4.948144656	15.86536423	14.33536469	8.45309955	9.266972906	13.31729137	14.50891929	2.395218418	1.893271462
POLE2_1	3.911200255	5.324362994	15.70196242	13.51521626	7.041801125	8.137035174	10.70146842	13.77877597	1.729834071	1.898597627
POLE2_2	4.281959734	4.492447621	15.15123542	12.37463932	6.862344995	7.71073903	12.33289602	13.74556016	2.695548595	2.30774475
POLE2_3	4.76724762	5.302382206	17.27660649	14.47199546	9.006058387	8.940506423	14.90996908	16.07767249	2.278373483	2.089273613
POLE3_1	4.05871928	4.307406356	15.86547582	14.08389086	7.369345788	7.673041502	13.80227415	15.41143291	2.356937755	2.155006775

POLE3_2	4.350770819	5.556651602	16.21218471	14.14844611	8.567572035	8.612137061	13.93869043	15.26105767	2.225090294	1.992268331
POLE3_3	4.463474364	4.881065731	15.78755662	13.76205423	7.919046954	8.534978589	14.42126801	15.22224553	2.117631687	2.221461599
POLE4_1	4.250250222	5.179949945	16.05761184	13.93492798	8.646373663	8.996169675	14.2918723	15.49072039	2.289609739	2.16380568
POLE4_2	4.203583788	4.609359523	17.20884099	13.73173182	8.284164222	8.694680875	14.46642374	15.18379211	2.320011701	2.370248198
POLE4_3	4.103565563	4.913910079	15.24091575	13.67414854	8.65679675	8.649365629	14.14675369	15.29752319	2.373757319	2.292665876
POLG_1	4.504529456	5.024635794	16.07528345	13.89986639	8.828294226	9.059076717	14.19249372	15.55583364	2.19194077	2.392965258
POLG_2	4.486580562	5.071883102	17.05550717	15.56254164	8.607868094	8.972231088	14.83591317	15.59590496	2.564599631	2.123748505
POLG_3	4.073551606	4.686363283	16.83335933	15.02678102	7.674746962	7.827473285	14.69259756	15.33641006	2.377130821	1.982573557
POLG2_1	5.241039803	5.722169338	17.37919755	15.58341001	9.168078886	9.401249685	15.18432967	16.06995638	2.462365588	2.281290394
POLG2_2	5.167876226	5.380124504	17.44865123	14.69148804	9.269627222	8.955149748	14.95370603	17.01997927	2.226136898	1.988044736
POLG2_3	4.933614967	5.490302455	17.21921826	15.90649453	8.918732211	9.067854443	14.58818371	15.72757165	1.774463426	1.935522621
POLH_1	4.272094541	4.773019358	14.85747393	14.22746454	8.562505765	8.343494868	13.59338135	14.71622202	2.140309872	2.249353688
POLH_2	4.27153214	5.110749128	16.70506337	14.6091716	8.802243005	8.901920808	14.08330504	15.51934304	2.382198006	2.227336123
POLH_3	4.357244266	5.144408389	15.44303988	14.1785632	8.746127873	9.218176069	14.49827529	15.50590157	2.189462747	1.90957074
POLI_1	4.150888436	4.906908665	15.46321208	13.81161491	8.257185023	8.711378839	13.95438687	15.32107119	2.232167523	2.35857237
POLI_2	4.102312561	5.017107148	17.13201857	14.96206292	8.37017365	8.790453461	13.92781768	15.63086331	2.177688149	2.09670287
POLI_3	4.303504027	4.884938029	15.30718031	14.32850766	8.206351138	8.396446257	14.29144885	15.50054455	1.973162917	2.38752182
POLK_1	3.890619592	4.923647277	16.35482832	14.0259265	8.520467971	8.574598094	13.84780368	15.71974759	2.212790909	2.052740662
POLK_2	4.470517162	5.017053527	15.06136798	14.30769741	8.559703681	9.176052009	13.75309838	15.33738671	2.263776628	1.821780289
POLK_3	4.366677089	4.543012816	16.26672905	14.34958803	8.256754173	8.456452682	13.57376505	15.01819955	2.290760174	2.263328273
POLL_1	4.524229503	7.023973639	19.07283677	16.83114506	8.837094695	9.629079016	15.31628842	16.86975457	2.191695531	2.246820619
POLL_2	4.071305348	4.509734513	15.90107192	13.74321881	7.365057554	7.517407267	13.67234026	15.06371209	2.396095536	1.986704711
POLL_3	4.430532533	4.917564771	15.72173799	14.5007933861	8.491153386	8.586252928	14.30921533	15.25891069	1.814817914	2.212447193
POLM_1_1	4.257313596	4.971430263	16.36817103	14.22212734	8.229618213	8.764506678	13.98620056	15.34016942	2.225210625	2.216900079
POLM_1_2	4.396375113	4.520193875	15.50395237	14.0878161	8.523054478	8.599533287	14.78026756	15.40935421	2.565521057	2.051981064
POLM_1_3	4.511963088	4.868661377	14.57249071	13.37953696	8.347456643	8.557697712	13.43649691	15.47598405	2.519285639	2.224112403
POLM_2_1	4.376989617	4.796331015	15.80917874	15.78572797	8.764063603	8.661289156	13.7911128	15.047217	2.439411654	2.12770753
POLM_2_2	4.309905317	5.122559299	17.09843593	13.66848179	8.796739874	8.864411597	14.25779264	16.00362077	2.085064806	2.057841084
POLM_2_3	4.357766661	4.698533862	16.31967902	14.19205994	8.219222668	8.722211438	14.12710602	14.79145586	2.134034076	2.273157539
POLN_1	4.538238533	5.438999856	17.18044954	15.30051909	8.121868486	8.256011787	15.39814149	15.95846412	2.325700266	2.335966298
POLN_2	4.44974539	5.606148397	15.73580612	14.6652103	9.565191334	8.934813283	15.01396877	17.09553131	2.307699486	2.226389295
POLN_3	3.896689556	4.566665524	16.30023077	14.63746462	8.443615554	8.885591286	15.61185027	16.55668123	2.353681262	2.41373351
POLQ_1	3.787417464	4.313139248	15.57647029	13.98702791	7.796867453	7.82973546	14.40364	15.63019555	2.298108934	2.35435043
POLQ_2	4.191465113	5.106375509	16.99503387	14.9110522	9.332580099	9.43652131	14.82183024	16.38097447	2.318034307	2.47664097
POLQ_3	4.444836578	4.797916247	15.89895829	13.39954529	8.417277603	8.585507869	14.4591784	15.90847101	2.558995149	1.928881427
POLR2I_1	5.569038191	5.954720205	19.48483038	18.47261004	10.32883082	10.52402814	15.51832283	17.03036462	2.948348962	2.871263627
POLR2I_2	5.93382847	6.344696877	21.06448985	18.32362374	9.426981049	9.596286652	15.66101854	16.67842563	2.506898594	2.306849471
POLR2I_3	4.003364172	5.362537764	14.96542185	13.7966222751	9.117842653	9.131510367	14.87399812	15.96528174	2.516477073	2.506760754
POLR3E_1	4.267454932	5.173887193	14.94288023	12.73486431	11.23406079	12.36755295	19.08840987	20.40880574	3.719219643	3.992382688
POLR3E_2	4.61907433	5.043676006	16.40335455	14.80724157	8.4444287	9.21468468	14.48991254	17.02045111	2.350612051	2.281793169
POLR3E_3	4.534494033	5.793629669	17.39323795	15.97547574	8.100074476	8.664091221	14.61973538	15.21184663	2.216568404	2.138299463
PPP4R1_1	4.296766294	4.67089355	16.35340901	14.01058037	8.367576392	8.447882649	13.77323163	15.24942315	2.251697439	2.117448106
PPP4R1_2	4.394722919	4.874407731	15.45676786	14.08490301	7.896272437	8.300763337	12.67199482	15.28951044	1.847010627	2.08612654
PPP4R1_3	4.757650212	5.083019602	15.56207887	13.94467917	8.5021114441	8.70797353	14.46543152	15.49975657	2.2402472	2.189793393
PRKDC_1	4.656510265	5.565339227	19.51500177	17.00721766	7.827691378	8.064022736	14.53181503	15.79205939	2.46291731	2.041787069
PRKDC_2	4.518295045	5.475420942	19.27621543	16.59777316	8.770791691	8.949845943	14.76760051	16.23188559	2.870366073	2.668812353

PRKDC_3	4.055667776	4.893084049	17.80901269	15.25352829	7.480956892	7.598001255	14.06358382	15.06318986	2.310322276	2.31254986			
PRPF19_1	6.35805115	8.523852385	19.37266207	17.11221754	7.449852528	8.328782639	12.64032662	14.51389826	2.129676441	1.962179			
PRPF19_2	4.141878999	6.926187759	18.70263395	17.01332087	7.731409264	8.442965485	13.51614734	15.3303293	2.417440104	2.272771271			
PRPF19_3	6.606727113	7.060041097	19.98858595	18.50814539	6.659019945	7.423851316	12.34848261	13.76819901	2.117941408	2.01691034			
PSIP1_1	4.525724934	4.738497696	15.97828896	13.83626078	8.490803256	8.768951941	14.44440466	15.6568471	2.256583691	1.740610531			
PSIP1_2	3.972018976	4.434315353	15.40627934	12.56666071	7.927606466	8.539382167	14.10966614	15.23228307	2.127376805	2.10746767			
PSIP1_3	3.934349216	4.533663321	15.43298669	12.58689347	8.205926879	8.717389539	14.20783052	14.86394054	2.162379843	2.253105331			
PTN_1	4.286120939	4.882620268	16.30849292	14.99121772	8.306894906	8.476311945	14.34785973	15.58345894	2.165702415	2.369224179			
PTN_2	4.638100611	4.944638253	16.23323466	13.16334591	7.753509761	9.07754443	14.39869533	16.17431889	2.334910658	2.098841563			
PTN_3	3.872373249	4.854657695	15.06819707	13.07015931	8.45797213	8.13693137	14.04747605	15.91478042	2.380173741	2.142670742			
RAD1_1	5.124074946	5.64785608	16.89339055	15.18086192	8.731115891	9.250245853	14.67416609	15.96708071	2.48375826	2.45312001			
RAD1_2	4.566380162	5.220640919	17.34872959	15.67864672	8.562402717	9.213729309	13.99398471	15.97692768	2.435272303	2.28812558			
RAD1_3	4.612093656	5.061667544	16.92884777	14.47162482	8.296539103	8.632593329	14.54141698	15.68794264	2.370148394	2.006167695			
RAD17_1	4.561158483	5.280591691	16.57395845	14.10451469	9.315035533	9.541009603	14.41482269	15.73051767	2.698703121	2.549535234			
RAD17_2	4.520352118	5.115205109	16.3358457 9	14.93191619	9.039494719	9.178402273	14.9093688	16.2968055	2.670407818	2.799204771			
RAD17_3	4.554664497	5.164607963	16.29166227	13.9863031 1	8.668947653	9.598784291	14.84568624	15.92267313	2.665017576	2.631158987			
RAD18_1	4.066790182	4.591136212	16.23181208	12.7936366	8.320268438	8.543560013	13.31111784	14.70384759	2.178153199	1.915426376			
RAD18_2	4.109109017	4.687919463	15.59802027	13.44523797	8.610763288	8.847568423	13.3471672	14.6375922	1.951670037	2.220455449			
RAD18_3	4.744131667	4.649769146	14.72169122	13.74312714	8.392061255	9.579086062	13.7075345	14.87862085	2.186583088	1.583034646			
RAD21_1	4.912288737	5.388832301	15.92680062	15.43525763	8.422124055	8.937647926	14.21003805	15.68986011	1.899575266	2.311769194			
RAD21_2	4.356129542	4.734827775	15.75056155	13.01591503	9.033362937	9.285171033	14.73819189	15.86417427	2.510297558	2.528765241			
RAD21_3	4.373251058	5.300474247	16.63273326	14.67732543	9.313013541	9.432315297	14.92951381	16.588233 86	2.318494927	2.148594137			
RAD23A_1	4.92377188	5.060198333	17.84539615	14.90275379	8.34990648	9.100150393	14.48524553	15.84353781	2.187558194	2.274480201			
RAD23A_2	4.517962545	4.946413072	16.36907999	14.30063323	8.272554701	8.690111253	13.76928662	15.72751148	2.052966763	2.139469713			
RAD23A_3	3.486108407	4.536719712	15.69813019	13.95976284	7.754121608	8.033558953	13.50137326	14.36417877	2.331661421	2.28968117			
RAD23B_1	4.847560225	5.320712929	15.75007449	14.55839356	8.686012745	8.676038123	13.85276267	15.56892191	2.121743296	1.924220205			
RAD23B_2	4.341275598	5.189964526	16.20449607	13.9650228	8.371553568	8.370788098	13.60068564	14.66927557	2.214909865	2.422921816			
RAD23B_3	4.187060007	5.652256073	16.37935266	13.03927582	8.571081263	9.291944797	14.63666884	15.9791395	2.529423268	2.064147349			
RAD50_1	4.166139364	5.314491358	16.37991456	14.86710516	7.804731462	7.997226645	12.93041706	14.36605023	1.972564193	2.250502344			
RAD50_2	4.504976427	4.990481488	16.6749246	15.96055395	7.858745935	8.569682715	13.84812702	14.77131914	2.057559578	1.854767367			
RAD50_3	4.24748274	4.940566332	16.40690439	15.05062528	8.199645206	8.315980948	14.4681566	15.70443164	1.976190476	2.116371605			
RAD51_1	4.436472897	4.373500646	15.82972999	14.14613207	8.513421022	8.887415507	13.77468124	15.78403397	2.665164108	2.291503208			
RAD51_2	3.947366031	4.908672526	16.52304609	14.04462614	8.765339836	8.663580493	13.42918062	15.32714512	2.502873969	1.985996181			
RAD51_3	3.793036374	4.503853999	15.88659606	14.84306592	8.222289323	8.332433391	13.78613032	15.24982298	2.065652225	2.109123365			
RAD51AP1_1	4.371300024	5.12237319	16.68409504	15.36384446	8.939983925	8.826660193	14.10338946	15.55791087	2.297557506	2.033340412			
RAD51AP1_2	4.301448179	4.866615652	14.90292957	14.11959526	8.778358965	8.493158868	14.30024865	15.3195577	2.328111952	2.249466092			
RAD51AP1_3	4.211462811	4.822443371	16.12861503	14.57365585	8.335743061	8.671820462	14.26782229	15.33257518	2.441558037	2.321129457			
RAD51B_1	4.573470044	4.662077419	15.13180628	13.51813504	8.433453425	8.377265176	13.68816376	15.15259487	2.041347 47	2.028654154			
RAD51B_2	4.411968504	4.970757999	15.97623896	14.10669379	8.375044094	8.769552363	13.76959782	15.59880773	1.840083105	2.107323899			
RAD51B_3	3.701242151	4.390960867	15.97255957	13.87519962	7.804421872	7.916762924	14.28839238	15.8334858	2.392502244	2.221402694			
RAD51C_1	3.416025748	4.044256068	14.54831741	12.61509947	7.464869473	7.779663632	13.07714142	14.39165743	2.271242851	2.305315284			
RAD51C_2	4.362511893	4.974160207	17.51240094	12.68820529	9.01621076	9.479086477	13.47578906	15.15161569	2.068179137	2.815688353			
RAD51C_3	4.114553521	4.932179813	14.84466765	13.66794065	8.377283824	8.963286432	13.54587596	14.23586974	2.032867121	2.438660124			
RAD51D_1	3.955697498	4.714238567	15.36687778	14.06233119	8.041412823	8.156630544	13.80791763	15.63613686	2.246476073	2.028437347			
RAD51D_2	4.267725266	4.71017175	16.0412851	14.35832956	8.240996445	8.489782356	13.57119184	15.30676393	2.551275638	2.641217317			
RAD51D_3	4.269172498	4.886402952	15.81552101	14.62033697	8.463162256	9.412328712	13.15560559	16.01315665	2.545615292	2.220594836			

RAD52_1_1	4.039299906	4.554501426	15.92766605	13.19202598	7.560191676	7.869057785	13.80109331	15.21787091	2.118963276	2.039400322
RAD52_1_2	4.275489044	5.123005724	15.9196836	13.91992705	8.485645578	8.585119127	13.59499285	15.15941658	2.407412803	2.343665212
RAD52_1_3	3.94093467	5.060122699	15.63917689	15.04676698	8.42801487	9.088671088	13.78412957	14.94056585	2.105528156	2.062139093
RAD52_2_1	4.702436451	4.857381177	16.03926374	14.9181853	9.08206643	9.419772681	15.44711155	17.33810533	2.349291128	2.468847716
RAD52_2_2	3.875209452	4.732425244	15.23526962	14.79466348	8.074606495	7.76302834	13.85172147	15.19760872	2.14614249	1.943921151
RAD52_2_3	4.043785457	5.076064452	16.14984097	13.81963502	8.692336821	8.795621134	14.84652205	15.87698827	2.340337009	2.01191051
RAD54B_1	4.251660056	4.701972213	15.42757331	13.08555803	8.464356966	8.407144967	14.37669161	15.37422449	2.170123413	2.446224573
RAD54B_2	4.121477991	5.558955956	16.485301	14.66434135	8.287212773	8.336227704	13.99849073	15.70624293	2.212578746	2.346560916
RAD54B_3	3.817814831	4.855457798	15.14285261	13.46579603	7.546431312	7.835306423	13.55634049	15.13454685	2.231782077	1.961664135
RAD54L_1	4.416815142	5.426102625	16.38230292	14.57874098	8.008107446	8.754368045	13.94009598	15.90670158	2.383490833	1.962598805
RAD54L_2	4.174020377	4.794256794	15.77773988	13.96475414	8.908665698	9.099507458	14.13519593	15.69292681	2.217260589	2.120468097
RAD54L_3	4.12262666	4.575843928	16.17092516	14.33532443	8.112555078	8.545643941	13.83684059	15.50009193	2.316887363	2.130221884
RAD54L2_1	4.023827784	4.819529427	14.97933394	13.8767165	8.459198472	8.539736997	13.88774236	15.13145866	2.279049586	2.235647662
RAD54L2_2	3.776206916	4.539252564	17.08209071	14.3207866	8.752677778	9.068024652	14.07600914	15.52435764	2.433655819	2.234298528
RAD54L2_3	4.57784391 3	4.933445609	16.41881257	14.40891377	8.476914574	8.920356016	14.29020683	15.76982265	2.147633448	2.220655622
RAD9A_1	5.125803393	5.272116225	15.94133794	15.67220317	8.657567046	8.326887561	13.77655246	14.40074734	2.411562725	2.196793628
RAD9A_2	4.166319011	4.558770192	16.66995609	13.97915336	8.477542903	8.917797125	14.02542373	15.40421706	2.267316457	2.089041926
RAD9A_3	4.867688427	4.963687631	15.93346195	14.62200796	9.305344436	9.319674104	14.52462615	15.49695004	3.155240897	2.584763245
RADX_1	4.189926293	5.047362556	16.81579144	14.20517093	8.368409184	9.006696544	14.26109688	15.12038947	2.164049415	2.301098846
RADX_2	4.33786422 8	5.134444373	16.34464397	14.50498046	8.544132051	8.537016484	14.30390865	15.70073762	2.532299428	2.127398746
RADX_3	3.925555582	4.926001067	16.67426679	13.80392155	8.638243452	8.649158029	14.05718333	15.33000296	2.243149605	1.969038171
RBBP4_1	3.869047619	4.944697463	16.19620232	14.15133542	8.865064043	8.958387919	14.1137593	16.0617219	2.468010158	2.192958441
RBBP4_2	7.255053885	8.046838407	20.79422071	16.33506578	8.941190044	8.675103879	14.86357239	16.16177338	2.619921321	2.614674026
RBBP4_3	3.901070305	4.490089442	15.27756934	14.19988994	7.346544224	7.172115886	13.58170112	14.04483326	2.29423691	2.243633519
RBBP7_1	4.507034647	5.283688523	16.50003346	14.46075173	8.101860066	8.368325468	13.37656537	14.57160221	2.113192304	1.979053397
RBBP7_2	4.106784692	4.960557745	15.18927467	14.1499368	8.190628546	8.605046158	13.80048423	14.69523666	1.917554223	2.244786172
RBBP7_3	4.689009664	5.292655558	17.25456621	14.90532842	8.070086213	8.418421471	13.90931058	15.00201386	2.213827549	1.996412698
RBBP8_1	4.273193509	5.01979997	15.72615506	13.76262479	8.852609078	9.147388313	14.23886037	15.36077987	2.355606961	2.164428754
RBBP8_2	4.524340606	5.174029597	16.17097857	14.03819058	8.532032755	8.907069687	14.10810704	15.39603307	2.007133942	2.082529497
RBBP8_3	4.467182431	5.460678298	16.01666061	14.30640124	8.566321163	9.100531736	14.44137289	15.25737	2.227364492	2.210434574
RBMX_1	5.460385439	7.133390504	16.1868212	15.45599152	7.481667357	7.507980039	11.71854967	15.04393017	2.255314557	2.333614824
RBMX_2	5.054355499	5.885201536	16.75184554	14.9571885	8.413834043	8.527757118	13.79106992	15.15295961	2.318807957	2.330324089
RBMX_3	4.769073866	5.284215101	16.51428973	14.96670297	8.755717825	8.95921927	14.42383516	15.5503373	2.34779259	2.405392983
RDM1_1	4.162459476	4.81493364	16.43009543	13.46409891	8.406083611	8.396403229	13.97024605	14.67962719	2.191364335	2.357665075
RDM1_2	3.925263551	4.343883587	14.27785155	13.21573418	7.271124137	7.934013007	13.70642531	14.4181681	2.376495395	2.227193153
RDM1_3	3.760049296	4.484918615	15.68006346	12.87446505	7.495374095	7.998370692	13.79466905	15.26998973	2.204955993	1.892646325
REC8_1	5.290157102	7.074307202	19.1251915 9	16.52931103	7.851312694	8.806605937	14.0050325	15.47113866	2.26756924	2.179620245
REC8_2	4.365382375	4.835532424	15.06019118	13.34560819	8.660755508	8.497891293	14.24519637	15.52039428	2.066284183	1.927127578
REC8_3	3.654004955	4.025515042	13.33930787	12.77089628	6.667745189	6.995559919	13.00784341	13.89900618	2.241785319	2.134321828
RECQL_1	4.265228973	5.149551619	15.86573316	14.16323166	8.63448 42	9.207954 74	13.71548083	15.7483018	2.199092396	2.170027021
RECQL_2	3.967953092	4.890932663	14.572366 8	12.83574936	8.536249938	8.948613484	13.48799266	14.66544274	2.379654228	1.98356516
RECQL_3	4.212627585	4.355082113	14.92278541	12.19627009	8.272644725	8.466095359	13.16007918	14.36314635	2.242790433	1.945277232
RECQL4_1	4.601007287	5.587627997	15.867524	15.01609077	8.070715253	8.186084662	13.3385225	14.61025297	2.409175247	2.161484956
RECQL4_2	4.208421214	4.650175536	15.61937162	13.85943961	8.576411343	8.698275221	14.45957843	15.31139376	2.057677775	2.07796391 6
RECQL4_3	4.462397188	5.106765803	16.42639663	14.65801696	7.393087712	7.49202395	13.1597531	15.28598624	2.199424252	2.233692614
RECQL5_1	4.10316136	4.830889909	14.72187651	14.26713704	8.430749012	8.364475384	13.56418667	15.19770863	2.2240854	2.079678405

RECQL5_2	4.205788404	5.305889387	15.23115896	14.3615364	8.341516916	8.305233921	13.84341549	14.50641799	2.345710532	2.170070511
RECQL5_3	4.45732799	5.311524616	16.0280765	13.94009274	8.320142525	8.596031449	13.62625074	15.50107818	2.554261065	1.936110168
REV1_1	4.136522408	5.051057986	16.45612319	14.91389039	8.100890038	8.966946198	13.75191121	14.76666115	2.093648184	2.202533329
REV1_2	4.131198286	4.579666515	14.89812838	13.796206231	8.05	8.630879094	14.25653343	15.27615809	2.285793723	1.994104774
REV1_3	4.401057513	5.018174673	16.85642593	14.77423282	8.371303241	8.638940289	14.28459016	16.35095161	2.130657134	2.172765067
REV3L_1	4.512259749	4.537246791	16.33371469	14.44604361	8.172605427	9.193592778	14.02775344	15.64116283	2.129412313	2.220015785
REV3L_2	4.159254456	4.813569582	16.2122402	13.69493373	8.827862835	9.235955056	13.83630203	15.61131023	2.180409106	2.316077289
REV3L_3	4.376449963	4.905131758	16.25850381	14.6648967	8.942042845	9.345620213	14.65379265	15.66432724	2.408824854	2.441352814
RFC1_1	5.880612272	7.052406471	17.39759882	15.49619896	11.91370847	12.22109356	17.13340492	18.62993852	2.378870777	2.037248851
RFC1_2	5.38217327	6.281405187	16.90853201	15.40407956	9.453383105	9.732197951	14.45449605	16.7697828	2.265310746	2.451706663
RFC1_3	6.747593135	8.311267843	18.95853899	15.81565855	13.78899175	13.91723561	17.92501758	18.77509493	2.658408394	2.262949098
RFC2_1	6.13191812	6.866003123	14.94597839	14.37354603	10.32835911	11.11409276	14.60316337	16.98879368	2.431640506	2.401078271
RFC2_2	6.575979701	7.925163946	17.45887383	15.43942015	10.58735271	11.02265106	14.06038439	16.65396694	2.73618367	2.522704338
RFC2_3	5.170846276	5.734416699	14.77347126	13.42801167	10.31620071	10.84808398	15.60454533	16.65131615	2.553831565	2.466064
RFC3_1	6.433801464	7.254726933	18.4412537	14.73667371	10.94497777	10.85307248	15.14458642	16.09409102	2.654264531	2.316098035
RFC3_2	6.149665023	6.849342049	16.64358853	14.23178036	10.51567145	10.42238421	14.95951759	16.74648622	2.530479494	2.177769165
RFC3_3	5.548583722	6.456750796	14.73327233	13.37442577	10.11642876	10.46656667	15.34512788	16.02338061	2.286626361	2.513485166
RFC4_1	6.624310941	6.214105956	15.86637262	15.72812052	10.03582682	10.67871563	14.59894532	16.33305637	2.741639015	1.869914831
RFC4_2	5.686255369	6.766941407	15.97470193	14.04792786	10.74229371	11.41998497	15.79461001	16.85112036	2.655711793	2.306674231
RFC4_3	5.866442001	5.759281369	16.4593483	13.73158347	11.06825314	11.1172052	15.82717954	16.91622342	2.401670727	2.071574219
RFC5_1	6.415552855	6.205090085	16.84631096	14.14541836	10.50716814	10.81746442	15.39395287	17.05686426	2.603691104	2.308894463
RFC5_2	5.644447633	6.025530582	15.40444588	13.81980531	9.207397684	10.21006987	15.14747295	16.7850846	2.274726628	1.99601621
RFC5_3	4.989313499	5.503844136	13.14129333	13.86841564	10.08981406	10.36584437	15.0916332	16.07833282	2.545666963	2.430737062
RFWD3_1	4.386909033	4.674212713	16.02733911	13.73095821	8.443716762	8.629800751	14.07298504	15.76776665	1.745842138	1.778830852
RFWD3_2	4.438915587	4.995854237	15.16789119	13.58389372	8.472062252	9.134697845	13.47333596	15.7360362	2.178154167	2.011509517
RFWD3_3	4.010502068	4.766270555	14.5257119	13.12536245	8.295467046	9.023608399	14.04015147	15.32476726	1.852851562	1.834507315
RHNO1_1	4.073960747	4.536107192	16.18418251	14.46676766	8.438407269	9.246483633	14.42119959	16.66062933	2.224210065	1.960057527
RHNO1_2	3.966318764	4.576386211	15.76276085	14.23981794	8.350017905	9.454509942	14.06836334	15.43049303	2.603351787	2.069325901
RHNO1_3	4.392783606	4.846940958	15.6895905	14.26919446	8.78282661	8.93077776	13.99892205	15.84046561	2.508144353	2.559243206
RIF1_1	3.994698342	4.699125021	15.07352271	12.05708614	7.740714156	8.099919726	13.37564054	14.56257011	2.148877143	2.17381928
RIF1_2	4.591328535	5.001766402	14.68767721	13.04416288	8.339676021	8.418554916	13.78564661	15.35548527	2.347255782	2.128089594
RIF1_3	3.539131406	4.316394788	13.3870435	12.52423975	7.662499901	7.539375161	12.84428774	14.11865649	2.331657265	2.306450373
RING1_1	3.856596664	4.459606186	15.22110607	14.0521127	8.773891653	8.768029951	13.96517845	14.92544141	2.203653012	1.956615681
RING1_2	4.434270505	5.055816091	16.23833176	14.53533561	8.780419165	8.733147328	14.55690151	15.52151517	2.105971995	2.045407838
RING1_3	4.123004512	4.978603966	15.99148876	13.70265669	8.147717257	8.213254444	13.56620156	14.98607889	2.194049866	1.892209317
RNASEH1_1	4.754691165	5.524887502	17.52333509	15.15618162	8.170192645	8.137055685	14.01262399	15.29076466	2.146351049	2.053718123
RNASEH1_2	3.568502645	4.19628616	14.60409124	13.00814063	7.719003352	7.99973719	13.2465046	14.4023852	2.150775869	2.299157401
RNASEH1_3	2.311915256	2.831693443	13.10870361	11.74364613	5.941642222	5.883789668	12.44836418	13.4897091	2.463162708	2.278694205
RNASEH2A_1	4.452525746	4.933942315	15.99466588	14.79691731	8.666151721	8.538144396	13.75376437	15.10598622	2.315844401	2.209683147
RNASEH2A_2	4.435986522	4.94321602	15.02717702	13.6405695	8.556610838	8.810023731	14.01364788	14.74433608	2.276414055	2.128577076
RNASEH2A_3	4.740973312	5.239606785	16.43211513	14.80366621	8.25037413	8.382194699	14.73803307	16.24850163	2.215414011	2.146541816
RNF168_1	4.101659697	4.54170761	15.58010033	13.9550395 3	8.690506112	8.834108161	14.2907533	15.89386947	2.103386343	1.915142934
RNF168_2	4.556018262	5.204677036	16.16002934	15.3648305 6	8.232231544	9.064249112	13.76948731	15.49367175	2.236982217	2.198506316
RNF168_3	4.063094556	4.895710435	16.23190145	14.90211861	8.155940798	8.513086374	14.60153343	15.81831095	2.331603667	2.274137199
RNF2_1	4.160125147	5.111450156	16.35001722	14.7493889 6	8.679224107	8.872078609	14.57306587	15.2947479	2.135780688	2.164712861
RNF2_2	4.089320755	4.993705691	16.9274481	15.31514016	8.510492388	8.35358453	14.91020041	16.28293376	1.967756597	2.369888226

RNF2_3	4.391771656	5.304811901	16.15856362	14.01440914	9.433925605	9.224551536	14.63695736	15.483198	2.173798436	2.052752619
RNF20_1	4.499890162	5.015538485	17.21736614	14.14628033	8.731181825	9.280429823	14.58796146	16.19409741	2.294266348	2.323121642
RNF20_2	4.366296244	4.779780205	16.41643758	14.1328	7.509012036	7.861015734	14.36000185	15.7560971	2.362948402	2.416109782
RNF20_3	6.059390048	5.961416857	17.50845424	16.10072093	13.95097201	14.43697099	21.18602325	23.04947263	5.767855724	6.147253759
RNF4_1	4.043084782	5.027160062	14.00639031	13.60316988	8.222343855	8.310008424	14.14047124	14.8744436	2.082800375	2.564315905
RNF4_2	5.765901943	6.719296131	19.04450919	17.4922852	11.68575162	12.60470543	18.41789796	22.21778283	3.305127837	3.347597744
RNF4_3	5.717395553	6.717418111	18.66921791	17.07110358	10.12134364	11.16259347	16.36719721	18.73944988	2.751074062	2.449839536
RNF40_1	5.768143635	5.510150277	18.49995423	15.62239142	13.77371616	13.45555958	20.66091994	22.47967621	4.748090721	5.032254088
RNF40_2	4.065613156	4.935667557	16.13694319	14.09412046	8.372930763	8.852952673	14.73709137	16.10562006	1.976685792	2.118772108
RNF40_3	4.28720967	5.041396453	15.01459001	14.42141771	8.984389952	8.762210541	13.60447657	15.22810622	2.263873521	2.374662103
RNF8_1	5.448181403	5.553617883	17.87353792	16.76415188	8.755567929	9.466396647	14.23526357	16.03480245	2.29603824	2.690339588
RNF8_2	4.456223243	5.176668164	17.03459698	16.01357984	8.278153881	8.738016138	13.96459553	15.59863764	3.117790242	2.741398217
RNF8_3	4.515231174	5.054618338	16.72640116	14.07114559	8.311705278	8.407262334	14.09093554	15.49223849	2.353354128	2.149579
RPA1_1	4.900389283	6.908315565	16.30237765	13.2342277	7.791208458	8.387248955	13.29232095	13.40368098	2.500854842	2.547431861
RPA1_2	4.536183405	6.09562717	15.235851115	13.32374901	8.05463463	8.012604563	13.5103348	14.20643867	2.405963303	2.247951464
RPA1_3	3.800475059	5.715952173	14.2820709	12.9867640	7.403612208	8.297684525	13.58291457	13.59831299	2.660880093	2.554522195
RPA2_1	4.199510737	5.203028464	13.03688756	13.01882831	6.969746258	7.432019554	11.46471523	12.91866029	2.427895191	1.909009813
RPA2_2	3.797468354	5.155658046	12.5057418	12.13849828	5.884560065	6.695608616	10.74691806	14.52830169	1.999470339	1.911969727
RPA2_3	5.295661572	4.823428079	14.7554586	13.50697053	7.068446922	7.773741567	12.20556745	12.38434164	1.631357496	1.63868368
RPA3_1	4.966120377	5.800560691	16.58471689	14.2852830	8.733636678	8.793133649	13.75635817	14.13447539	2.511043616	2.2868307
RPA3_2	4.06562073	5.204398982	16.96659303	15.39544425	8.131543374	9.031417957	13.65038532	14.62137725	2.480137849	2.217907592
RPA3_3	4.978183261	5.106281914	13.63676692	11.49425287	6.726645004	7.507367768	11.86307681	13.44930879	2.667872902	2.702016193
RPA4_1	4.237288136	5.139698505	15.85751026	14.64951011	8.489647464	8.625049648	14.21008757	15.617105	2.214829788	2.288403206
RPA4_2	4.408678298	4.956801816	16.09812212	14.82775942	8.650433574	8.936345102	14.1813668	15.28880503	2.132527766	1.902566656
RPA4_3	4.301909431	4.955587033	16.20311515	15.13685527	8.631679534	8.905542573	14.46949609	14.82560361	2.084513277	2.197755529
RRM2B_1	3.606589231	3.871189706	13.39710787	12.7491137	8.235323874	8.211683774	13.59405628	14.89797591	2.447684721	1.862715387
RRM2B_2	4.200634058	4.770675234	16.22879037	13.93756396	8.621841136	8.724809075	14.03919619	15.27831658	2.03066219	2.127616338
RRM2B_3	3.413781976	3.878664439	13.53852332	12.37886065	8.112585431	8.06249985	14.04465881	15.22135378	2.145409192	2.336886453
RTEL1_1	4.508331904	5.161681534	15.14760173	13.92336139	8.878315564	8.640348983	13.80469976	15.81937758	2.235462307	2.271941949
RTEL1_2	3.933158272	5.094933473	16.89677736	13.71964124	8.671578427	9.137439976	13.79741644	16.04829427	2.136547494	1.958014709
RTEL1_3	4.638926374	4.745122429	15.27322589	14.69175482	8.707704072	8.898623025	13.87303472	15.46962004	2.295631017	2.122161042
RUVBL1_1	4.035551285	4.610921148	15.43768275	13.97685185	8.276571084	8.706114506	13.8638144	15.19114391	2.46758425	2.709186791
RUVBL1_2	5.160886671	4.785362421	15.83193494	14.25996305	7.845426386	8.373680219	12.84163596	13.82234432	2.322609573	2.291962895
RUVBL1_3	4.011627907	6.16567728	16.07600165	13.64949446	6.868581147	6.981685293	11.82767279	12.42132234	1.968775284	1.937913299
RUVBL2_1	4.403665636	5.150602136	16.65172977	14.4544176	8.309075059	8.615495843	14.32150371	15.33989347	2.26972078	2.17575188
RUVBL2_2	5.516492851	5.176152525	17.04089816	15.53736511	9.758083507	9.836991528	16.01577124	16.99779523	3.309056598	3.238476863
RUVBL2_3	4.677474653	5.608599431	16.58044675	15.2285824	9.27164287	9.263105718	15.10202758	16.87115182	2.657866279	2.846204894
SAE1_1	7.736048664	9.034581575	22.94166716	20.36829851	10.83089419	10.48079742	17.15889898	16.4942791	3.352075543	3.78711528
SAE1_2	8.362221336	9.899212234	23.62546885	19.02854056	10.252995	9.781098617	16.50746819	16.04885657	3.715785292	3.815374306
SAE1_3	7.344776006	8.311755942	21.54812498	19.18123882	10.49287109	10.03912425	17.46878746	18.15080497	3.47510534	3.532809954
SCMH1_1_1	4.524031348	5.142536686	15.22187982	13.93207004	8.514115558	8.454301647	14.00883699	15.35855424	2.427374083	1.874860374
SCMH1_1_2	4.26625501	4.839260532	15.60199027	13.9320700	8.514115558	8.454301647	14.00883699	15.35855424	2.427374083	1.874860374
SCMH1_1_3	4.740755527	5.142052578	15.77588044	15.24794016	8.719217378	8.778009979	14.71042589	15.26223659	2.487045351	1.977425678
SCMH1_2_1	4.473363156	5.095366471	15.30409653	14.44584311	8.445101022	8.498507946	13.99957769	15.30476429	2.095626488	2.468128261
SCMH1_2_2	3.956096633	4.92869017	16.18725508	14.20717799	8.194583384	8.651812604	13.74720674	16.30627625	2.475223669	2.224380316
SCMH1_2_3	4.032960199	5.212035504	16.31909293	14.40951719	8.46880597	8.983768019	14.62077592	15.5227874	2.191022068	2.051567944

SEM1_1	3.899039367	2.73287591	9.942543093	8.619282257	6.786091764	8.28332745	12.91654282	14.35893441	2.151792553	2.271339994				
SEM1_2	1.811125486	4.768051465	11.27415016	8.710257948	7.713659074	9.192789721	11.93262853	14.39507598	2.535606918	2.648185111				
SEM1_3	4.067985511	4.531867708	10.89978453	9.163172369	8.524004308	9.570595446	13.94418437	15.91021519	2.623382321	2.862349385				
SETD2_1	4.314307478	5.158470973	16.56013692	15.06797046	8.133187473	8.602780182	13.39737938	15.0897135	2.186387221	2.141639385				
SETD2_2	3.717820201	4.362517902	14.60915429	12.66910896	8.634422162	8.602250713	14.30536556	15.69725023	2.163307041	2.106424893				
SETD2_3	4.510503913	5.363806957	16.80931666	15.08513478	8.232402105	9.135557249	14.49601873	15.01486458	2.302070521	2.056515401				
SETMAR_1	4.110770035	4.560563577	15.42428212	14.56238957	8.101751695	8.219641686	14.72635167	15.38640075	2.221015823	2.276961824				
SETMAR_2	4.097516696	5.174694242	14.85793006	12.99162311	8.916396647	8.783677102	13.86010143	15.60321892	2.231432361	1.906733353				
SETMAR_3	4.218452221	4.934430665	16.10914433	13.84355876	8.64525966	8.560169699	14.56234536	15.58235642	2.573587	2.179599912				
SETX_1	4.669887279	5.011960561	17.17460162	15.66295099	7.847118903	8.349799868	14.27503666	15.885146281	2.305792716	2.493501783				
SETX_2	4.813781328	5.758684455	17.92120369	16.97484628	8.891658622	9.150885047	15.07731664	16.04441278	2.679345209	2.404531381				
SETX_3	4.451698197	5.387755711	17.02028034	14.89724248	9.105366386	8.861257919	14.78035284	15.96892008	2.306163581	2.560291633				
SGO1_1	5.384937236	4.98498182	16.28358084	15.23908046	7.641403421	8.152631982	12.8964539	13.93709563	1.848549691	2.663656885				
SGO1_2	4.646602747	4.674582478	15.44588822	13.77673875	7.45937131	8.87114625	12.52045365	14.44801415	2.296859933	1.359206702				
SGO1_3	3.944547663	5.241113678	16.36123251	15.47736876	7.343179555	8.368851904	12.57302383	12.36910995	1.687592131	1.817047493				
SGO2_1	4.339501397	4.602527551	16.02884803	13.97145658	8.802763943	8.526682663	14.23831606	15.00937462	2.294757053	2.271606098				
SGO2_2	3.860383136	4.800424967	16.4785142	15.26020361	7.565863264	7.336196706	14.73135901	15.77702282	2.426620666	2.340736566				
SGO2_3	4.121106382	4.70777611	15.36493027	14.08415714	8.085591798	8.673498467	13.79161146	15.37713301	2.125630604	1.9566655				
SHPRH_1	4.084784711	5.102159411	16.55218896	14.47058905	7.815223826	8.504234962	14.17671468	15.45361431	2.397421905	2.412464785				
SHPRH_2	4.391392784	4.997752334	16.27728676	14.37304016	9.01113264	9.050932556	14.4139883	16.01735627	2.156047965	2.465454845				
SHPRH_3	4.694229371	5.230785135	16.33346013	15.58603368	8.717831117	8.966851702	13.80766034	15.35095805	2.298944538	2.081742613				
SIRT6_1	3.865395305	4.736328125	14.87373365	12.52511351	7.589053413	7.957674676	13.07240522	14.17241501	2.112593597	1.938145223				
SIRT6_2	4.418967275	4.698163108	16.276824	14.26379209	8.293091264	8.111282866	13.16532	14.52338528	2.327223471	2.209420386				
SIRT6_3	4.27751551	4.507165918	15.05346504	14.1575437	8.922996182	8.398526337	14.74601036	15.9493462	2.187956027	2.199341929				
SKIV2L_1	5.518763797	6.786551994	23.48136823	13.86943121	9.413306607	10.27667984	16.13355667	15.43398979	3.314457028	1.241743725				
SKIV2L_2	4.155088643	5.014572351	15.32331291	13.13957124	8.56485271	9.242409615	14.117435	15.23539253	2.431788032	2.103507715				
SKIV2L_3	4.408440467	4.812486114	15.73375766	14.07798872	8.71042753	8.663278473	14.39537733	15.32990071	2.154408758	2.064300536				
SLX1A_1	3.685774667	4.867674672	14.68026663	14.47292647	8.372486141	8.537433208	13.74741201	15.04634958	2.122144579	2.146582231				
SLX1A_3	4.184460352	4.836775483	14.84060213	12.93175998	8.595891117	8.705111533	14.66297394	15.87469354	2.088179703	2.280940814				
SLX1A_2 SLX1B_2	4.173821976	4.738669334	15.4745 59	13.92464714	8.403831832	8.469561167	13.93024039	15.09464808	2.307095523	2.224400695				
SLX1B_1	3.864848179	4.665893328	15.81574643	14.08845082	8.217229056	8.196154786	14.28374997	15.53373935	2.058524757	2.076607712				
SLX1B_3	3.846327746	4.852910613	16.7631583	13.90463825	8.167120816	8.062996368	13.88167366	14.84279671	2.350291516	1.997639919				
SLX4_1	3.916537431	4.452746365	15.16576412	13.12415459	8.298527681	8.993112658	14.39761096	15.12559114	2.153339323	2.269966736				
SLX4_2	3.431507985	4.123335801	14.17599363	12.56406805	7.996007541	8.727637519	13.79242261	15.11075869	2.523691407	2.347626087				
SLX4_3	3.791537312	4.263639933	13.44933951	11.31590649	8.400039186	8.867793536	14.22483464	16.08365938	2.184564259	2.323657058				
SMARCA2_1	4.338895025	5.301257044	15.46065991	14.56186876	8.870789079	9.326811014	14.28537121	15.91454996	2.198775982	2.230379638				
SMARCA2_2	4.064849624	4.911516568	15.5609427	13.90628549	8.665823464	8.951798011	13.6242433	15.16947786	2.220150669	2.119383702				
SMARCA2_3	4.391348721	5.158001675	15.78850814	13.7666647	8.267144402	9.224775178	13.8886526	15.46728643	2.170415195	2.283317164				
SMARCA4_1	4.466193961	5.290424332	15.93530274	14.60190258	8.642840597	8.5995086	14.1739488	15.7474868	2.305398024	2.084933436				
SMARCA4_2	4.412370285	4.731294745	15.84764858	14.60819583	8.38351394	8.875564924	14.70790716	16.17020646	2.194879426	2.468256215				
SMARCA4_3	4.402837242	4.327675602	15.37455933	14.16068166	8.610124565	8.88332603	14.50840985	16.23912147	2.582617523	2.161937428				
SMARCA5_1	6.767337046	7.961868303	21.19639644	19.20236337	8.132784263	8.331230252	13.62602963	14.52635403	2.168187754	2.348199522				
SMARCA5_2	6.165326281	6.978325401	20.84068452	18.16550943	8.064630983	8.362083836	13.57697471	14.81308605	2.432469736	2.348951605				
SMARCA5_3	6.191558158	7.419150285	20.10657308	17.98928328	8.593933143	8.426504439	13.75954996	15.76387196	2.186669613	2.03347026				
SMARCAD1_1	3.923003321	4.261646674	15.01191104	13.45681771	7.734115127	7.235023663	12.56067269	13.97616003	2.040945613	1.996156417				
SMARCAD1_2	3.703020971	4.512294336	15.89930216	14.20538492	8.167365906	8.135863913	13.95105907	15.13129478	2.061205523	2.262892013				

SMARCAD1_3	4.172830731	4.24145036874	14.886386101	13.89318126	7.79337106	7.350924312	12.34431365	13.81454459	1.771490186	1.633588697
SMARCAL1_1	4.009719174	4.562829844	14.88066982	13.76890875	7.688560494	8.104130865	13.5410925	14.84506448	2.41222605	2.523041339
SMARCAL1_2	4.30357691	4.760941096	14.64187177	13.55128235	8.746308288	8.801097871	13.59280254	15.236289	2.109169688	2.256443676
SMARCAL1_3	3.849678074	4.611281398	14.77847474	12.94874363	8.32457101	8.714855572	13.62769629	15.40267668	2.040386505	1.762053329
SMARCB1_1_1	4.439186235	4.751226456	15.32519877	13.76337656	8.388207039	8.209606157	14.01231678	15.66931857	2.361507909	2.263901147
SMARCB1_1_2	3.775617411	4.794309977	16.64222151	14.97103445	7.872412153	8.550883396	13.78794392	15.35761741	2.178199699	2.034247288
SMARCB1_1_3	4.587281031	4.5957901	15.71211315	15.09975583	9.073280173	9.582679244	15.1263316	16.46558068	2.535230674	2.223597104
SMARCB1_2_1	3.931720209	4.771628157	16.12814595	14.16734382	8.327742163	8.480921018	13.80080636	15.56321685	1.904629866	2.10665945
SMARCB1_2_2	3.840948796	4.842657669	16.30781631	13.92381511	8.449751184	8.869662671	14.04559166	15.69666062	2.132221829	2.042059633
SMARCB1_2_3	4.538714833	5.187797316	15.75946237	14.1594104	9.102034786	9.095779221	14.07311368	15.61406179	2.293575509	1.898746219
SMARCC1_1	4.241120368	4.432966972	16.29124756	14.60989377	8.897016012	8.915768157	14.56491614	16.23024167	2.588118971	2.158465852
SMARCC1_2	3.624647604	4.541463975	14.79883406	14.153733851	8.165586706	8.126002595	14.65104283	15.96706481	2.495981629	2.564739563
SMARCC1_3	4.455767563	4.378478664	15.87524445	13.62376238	8.328606896	8.815754692	13.97390995	16.31320709	2.368518225	1.890116931
SMARCC2_1	4.328668517	4.841892044	15.73854802	14.51503629	7.91163343	8.288457954	13.75899069	15.49711501	2.185014345	2.296242711
SMARCC2_2	4.533248373	4.89238058	15.42959111	14.69648562	9.018512342	8.771752071	14.45497854	15.27039075	2.331188185	2.112490484
SMARCC2_3	4.463662954	4.743645432	16.62923003	14.2947328	8.493564071	9.016227367	14.3516458	15.60451789	2.209612592	2.062641621
SMARCD1_1	4.402959209	5.247755535	16.24239357	14.82877788	9.868075202	10.04503756	15.69538556	17.53050514	2.543752544	2.435167276
SMARCD1_2	4.269459568	5.039509311	15.71201837	14.86224135	8.351435601	9.153863727	15.9661773	16.99533163	2.854628773	2.562198272
SMARCD1_3	4.561944468	4.840129376	16.04015251	14.53444577	9.122729217	9.032790409	14.59776503	16.05846614	2.290739044	2.579301898
SMARCD2_1	4.024053463	4.608785636	15.60426434	13.9581855	8.437386535	8.996970812	14.49158914	15.10662505	2.402312042	1.998404922
SMARCD2_2	4.559007666	5.263570493	15.79151024	14.8993901	8.131181246	8.595870164	13.97690224	15.57346222	2.344059759	2.394056136
SMARCD2_3	4.592103361	4.953551623	15.29736542	13.72217968	8.599697255	8.571854241	13.86370015	15.04221693	2.179459634	1.897346256
SMARCE1_1	3.897641775	4.298871637	17.22414775	15.56423334	8.72544665	8.629260981	15.78545349	16.82414186	2.748356682	2.640105456
SMARCE1_2	4.60621769	5.104597634	16.28200537	15.07886348	8.309910673	8.649169546	13.84959047	14.99380156	2.082009896	1.929566732
SMARCE1_3	4.828737919	4.986288804	17.82510507	15.85460871	10.05263444	10.30190678	16.43426874	17.39190181	2.470340183	2.792835396
SMC1A_1	10.93298615	11.51755452	29.4911121	25.94337259	9.530254815	9.539242121	15.09968126	16.08927897	2.940133076	2.763639188
SMC1A_2	5.811096704	6.903669224	20.60161691	16.75745385	9.893679465	9.837796415	15.02682617	16.45116216	2.779165733	2.627116152
SMC1A_3	6.407197873	6.450231283	19.58683045	17.08579742	9.699016765	9.30050165	15.49541662	16.52711436	2.718965688	2.75444412
SMC1B_1	4.16539961	5.017550991	15.65734288	13.72140732	8.537833781	8.822379536	14.2209521	15.40220642	2.129260597	2.056557394
SMC1B_2	4.173835566	5.293773254	17.15787659	14.15026614	8.687222269	8.936384121	14.26492663	15.45550127	2.116499664	1.961764706
SMC1B_3	4.502185076	5.093256234	15.76299757	15.00485401	8.812153368	9.109341208	14.02971133	15.76283066	2.316928061	2.176209102
SMC2_1	4.689279525	5.20327282	15.82197767	14.29774628	8.632660355	8.971449177	14.0438062	15.30691081	2.092703982	2.118528891
SMC2_2	4.375341589	5.046952076	17.24244144	14.62586616	8.272338126	8.32039517	13.85801846	15.23243751	1.977046443	2.136560421
SMC2_3	5.106026786	5.598626975	17.06460362	15.14537902	8.024831165	8.65464127	13.391632	15.38606962	2.315440296	2.183606797
SMC3_1	5.068395491	5.712140309	17.17676677	14.34282333	9.351435668	9.431470538	15.01230216	16.13929687	2.410602079	1.852994983
SMC3_2	3.926752201	4.581423131	15.12011266	13.36229545	9.489284185	9.987730061	15.58159532	15.52528012	2.575943789	2.755073732
SMC3_3	6.139653253	7.056823528	20.58954423	16.65669771	9.418384315	9.407674962	15.46885069	16.65601572	2.605380797	2.47317761
SMC4_1_1	4.253962975	4.88565614	15.77710916	14.2319699	8.542835658	8.730862101	13.55334451	14.95035545	2.057529205	2.06602222
SMC4_1_2	4.412751678	4.961201026	15.33757748	12.82790698	7.215113551	7.705748282	12.8444138	15.10303067	2.406357982	2.54055288
SMC4_1_3	4.407151227	4.960894365	15.88504739	13.97590361	8.030335089	8.756737462	14.59200566	15.77774623	2.008688985	2.251117562
SMC4_2_1	4.452666712	4.92830635	16.38471623	14.65596788	8.537675975	8.730396703	14.58607067	15.79203689	2.262867899	2.106131301
SMC4_2_2	4.456364971	4.969286613	16.39065456	13.95115579	8.342790617	8.623419999	14.16222742	15.53024348	2.214345244	2.077150093
SMC4_2_3	4.233000844	4.782419702	16.23268411	14.45416735	8.294002533	8.576269156	13.93927352	15.41186241	2.196242778	2.141432761
SMG1_1	4.345643443	4.421497726	15.76963774	12.97986732	8.863602626	9.005180363	14.16440563	15.70004135	1.963455962	2.145892966
SMG1_2	4.451287427	4.850348503	15.7631593	14.2273369	8.649251138	8.711792702	13.99645675	15.66779335	2.436142542	2.024993043
SMG1_3	4.443605401	4.97799176	16.34600409	14.33422099	8.569512063	9.154548149	14.47661123	15.66037263	2.273964017	1.936306259

SMUG1_1	3.899984694	4.669895676	15.381102781	14.11512512	6.782085294	7.0008823451	13.21433261	14.94070914	2.405467975	2.358652199
SMUG1_2	4.286041476	4.761654553	15.223958451	14.06271733	8.15900684	8.3218981011	13.92767209	15.27461362	2.386926673	2.074112088
SMUG1_3	4.277398635	5.022182128	16.504801041	14.38125483	8.37517055	8.4592194511	14.25341816	15.58640017	2.210932689	2.092856544
SNW1_1_1	5.456086462	6.424870466	20.472171921	16.84711396	9.77938243911	9.8087726981	14.90066325	17.20546098	2.862096714	2.682055966
SNW1_1_2	5.286065286	5.690418951	18.334316351	17.6890595	9.16286783331	9.4273764341	15.97944955	17.01704192	2.590874129	3.109362338
SNW1_1_3	6.383489321	4.704882024	19.502124351	15.86023211	9.15537671611	9.6093117411	14.69979926	16.02355881	2.8633261211	2.146736938
SNW1_2_1	4.407302455	5.260733762	16.37023661	14.29597311	8.37642083911	8.4155018241	14.15744953	15.71683128	2.241851684	2.046054732
SNW1_2_2	4.450582254	5.325811541	16.400927351	14.69033811	8.88822164911	9.3143930141	14.25548461	15.61600136	2.250694936	2.194540071
SNW1_2_3	4.333209477	4.809973575	16.463011271	13.91699025	8.71576395811	8.8996690221	14.13397725	15.16981348	2.082302757	2.354448273
SPO11_1	4.257807543	4.938289584	16.27940264	14.96435282	8.73042433311	9.07538709	15.02882318	16.63730221	2.318879164	2.075637835
SPO11_2	3.887809616	4.578131207	15.616220721	13.77274431	8.36507883911	8.5795996191	14.10007249	15.1538167	1.965444276	2.142667343
SPO11_3	4.739647713	5.103725041	16.204943121	14.20667042	8.83710095111	8.921436661	14.5626501	15.60033739	2.334085742	1.930955994
SPRTN_1	4.003875296	4.896415467	15.485339916	14.0709013	8.75119014511	8.5531122911	13.84142265	15.34390164	2.404831412	2.011926391
SPRTN_2	4.304680468	4.772021572	15.565225171	14.78439425	7.50758628511	7.5190468151	13.70530734	15.26218093	2.029586754	2.111407062
SPRTN_3	4.354217196	4.683514467	16.360831171	13.84469791	8.77277015711	8.6979241291	14.32882435	15.50904942	2.199505337	1.861961196
SRCAP_1	4.349110074	5.116959064	15.316833111	14.90553663	8.26638280711	8.5486513891	14.43660469	15.41572208	1.969414549	2.143571918
SRCAP_2	5.061579894	5.981804093	18.108894951	16.16831576	8.63118810311	9.0593787911	14.04869551	15.62456656	2.236216425	2.285988702
SRCAP_3	4.332957729	5.342604063	15.938003741	15.01370629	8.19469415911	8.1474194571	13.46093839	14.57956867	2.250219631	2.001046701
SRSF1_1	5.551354477	6.511761331	16.128965031	14.34611625	10.5617278511	9.7678652711	16.75625411	17.45134614	3.123432203	2.764473574
SRSF1_2	4.787368359	5.451837324	14.751021311	12.89734484	11.2403411411	11.093204407	17.56649057	19.49386161	3.581334242	3.080497312
SRSF1_3	5.899330684	5.747449594	15.139310511	13.87640828	11.0048481611	10.482997621	16.01996525	17.53155654	3.154780987	2.678840405
SS18_1	4.074861586	4.367540216	14.132962611	12.30150378	7.51558612211	7.7397653111	11.67028871	13.26650642	1.888686088	1.740747765
SS18_2	3.695356512	4.427059616	15.373229631	12.74291564	7.98913316911	8.1800156331	13.75065327	14.77889978	1.888127547	1.866688137
SS18_3	3.656920767	3.963628789	14.793142261	12.73628801	7.07082551611	7.5135905321	12.18649094	13.19437942	1.901312811	1.456305561
SSBP1_1	5.803607431	6.366574397	18.352706491	16.05229447	8.88763410511	8.8982433251	14.28767311	15.95355778	2.194109916	2.155880443
SSBP1_2	4.741275306	5.546387097	17.300434921	15.89366014	8.48238912681	8.4275327411	14.44585718	15.49123649	2.164603794	2.403661461
SSBP1_3	4.909347986	5.671845411	18.377329191	16.80888528	7.597930681	7.7443863211	14.13583255	15.52985763	2.275283415	2.230262335
SSRP1_1	6.172957477	6.78709441	18.247363881	16.12083946	8.93733829661	9.5402387371	14.37259567	15.68908166	2.196428433	2.336299777
SSRP1_2	5.640441291	7.219481908	19.481347941	17.91110275	9.20734753411	9.509311266	13.07979677	14.07577554	1.922052432	2.049564921
SSRP1_3	4.928021842	5.16907398	17.393725461	14.82348703	9.06475405911	9.8260406941	14.97353014	16.08782659	2.17362724	2.309215148
STAG1_1	4.348279893	5.039370079	16.52751656	14.5141635	8.78072309411	8.8260462121	14.54359539	15.80392695	2.054820466	2.035253126
STAG1_2	4.407220686	6.068096814	18.161404461	15.39176654	8.63155722911	9.0709321341	14.23167464	15.82277495	2.431492918	2.410293066
STAG1_3	4.042884633	4.802288667	16.633402931	13.29270285	8.94574789111	9.3455031181	15.36025245	17.04628514	2.657397935	1.958067793
STAG2_1_1	3.782584687	4.776613547	15.032707681	13.80595255	9.20498053611	9.5346036541	14.8978177	16.03278883	2.543075309	2.287227218
STAG2_1_2	3.457147285	4.577609443	14.570837621	13.41038166	8.55600570511	9.2270000191	14.8474092	16.09938228	2.392428715	1.995888983
STAG2_1_3	3.700847368	3.934946112	14.08182841	13.09069432	8.78080899511	9.3023021611	14.67914105	15.51976306	2.16948712	2.171750663
STAG2_2_1	4.175174389	5.05036543	15.146443111	14.01441973	8.08322067511	8.8056167981	13.66495791	15.43490095	2.190655998	2.142918288
STAG2_2_2	4.579005152	5.072650321	15.707099831	14.20040076	8.89612702911	9.4711883881	14.26086578	15.28857747	2.176341033	2.122816567
STAG2_2_3	4.577620522	4.798063167	16.815284911	14.62732471	8.44163375411	9.4156712781	14.77642088	15.66376099	2.303113854	1.848847368
STAG3_1	4.602612395	5.160592255	16.776940231	14.99453666	8.87919953611	8.3168229651	15.05491015	16.19628947	1.899276518	1.744497657
STAG3_2	6.010648063	6.650605203	19.432165671	17.84186327	8.456433191	8.3370486981	13.98017389	15.07662412	2.340955347	2.136845833
STAG3_3	4.473583158	4.955932792	16.235882531	13.91870305	7.93665111111	8.476206951	13.75118443	15.59657197	2.089670318	2.197813574
SUMO1_1	4.127949468	5.082479907	15.726295541	14.60250857	8.34990966771	9.0755056741	14.52399345	15.83375367	2.424044153	2.197874641
SUMO1_2	4.195302332	4.963544899	15.262842151	14.12654605	9.03115460411	8.7295966931	14.74216068	16.02333668	2.146464194	2.147824462
SUMO1_3	4.720581943	5.196087107	16.384513211	13.9489436	8.15642210771	8.0115604111	13.45536154	14.93181213	2.11517165	2.162829255
SUMO2_1	3.301640201	3.849723926	14.023522431	12.49174655	8.45846466921	8.9545634851	14.5174622	15.42419732	2.253503042	2.170346839

SUMO2_2	4.536352655	5.562558853	17.36357998	14.26402539	8.276944022	8.476891232	14.30017946	15.57959295	2.463286333	2.276671771
SUMO2_3	3.07253263	3.419704183	14.24267304	12.68526591	6.158971512	6.243206201	12.1506886	12.511127	2.45370332	2.297794118
SUMO3_1	4.110801421	4.886794464	15.86387053	13.47206607	8.902780049	9.016668831	14.68163061	15.74707231	2.088494262	2.566288631
SUMO3_2	3.839861796	4.394430113	14.8604461 5	13.55972894	7.837278423	7.848863745	14.03967859	15.22996404	2.225367134	2.147953674
SUMO3_3	4.470554362	5.065261645	15.82386238	14.02965772	8.099604866	8.499349988	14.85546993	15.06493647	2.253872996	2.035705457
SUPT16H_1	4.60364719	5.274524526	17.52771711	15.83640253	8.875615256	9.069814141	14.52050917	16.27216113	2.175338321	2.272787546
SUPT16H_2	4.776640266	5.725982149	17.4128701 1	15.85615039	8.821866905	9.164573757	14.44592442	16.453277	2.288087081	2.251170364
SUPT16H_3	5.213550678	5.99190165	18.44225293	17.07229614	9.149873934	9.256592514	14.5848929	15.5378198	2.010080255	2.120597711
SUZ12_1	4.558412113	4.720401145	16.61713229	14.21992405	9.529638124	9.610998673	15.13841106	16.37265886	2.411872401	2.338714412
SUZ12_2	3.988591981	4.74827912	14.64155479	14.11992936	7.993269379	8.979024514	14.33585843	15.75111937	2.323361227	2.221920945
SUZ12_3	3.788808917	4.778424667	15.14864462	13.63720083	8.765413063	9.309344576	14.83834668	15.68709491	2.067536702	2.063429961
SYCP1_1	4.426370713	5.081847704	15.40516104	14.33507495	8.794720007	8.7201825	13.84185466	15.87128662	2.273134864	2.073700712
SYCP1_2	4.395878345	4.887566523	15.32620155	14.05099351	8.639479015	9.007157951	14.09860811	15.53536024	2.218449493	2.032987035
SYCP1_3	4.648743 27	5.20823699	16.57410517	14.31991471	8.516932483	8.735490075	13.94746971	15.60146698	2.287333686	2.128253279
SYCP2_1	4.06418666	4.625380365	16.54533538	13.67656163	7.756279247	8.329852673	14.13014922	15.42983522	2.213614877	2.024540564
SYCP2_2	4.297109529	5.511273214	16.08256061	14.88391325	8.575102038	8.55097 93	14.26786778	15.63938062	2.154387404	2.009048903
SYCP2_3	4.507649501	5.474562035	16.13786012	14.46625733	8.659476276	9.110549772	14.92585555	15.55295944	2.282375496	2.157371901
SYCP3_1	4.332096983	5.35773031	16.82722 05	14.88245642	7.990362827	7.891547049	14.39488522	14.51350112	2.1856293 38	2.073015214
SYCP3_2	3.830210579	4.968731034	15.64839614	13.79770674	8.678321585	9.308555733	14.94264144	14.86709206	2.179983574	1.680512879
SYCP3_3	4.439062669	4.81021082	15.97747556	14.20560461	8.520606367	8.749972406	14.79362745	15.82496006	2.403767538	2.273882923
TAF15_1	3.952932049	4.840496698	15.5447699	13.87519195	8.399303995	8.618718252	14.11298552	15.28086945	2.651399331	2.155888017
TAF15_2	4.219906867	4.823523698	15.35247464	13.43696295	8.230697358	8.630039868	14.04753211	15.26679619	2.063851042	2.024256916
TAF15_3	4.545205629	5.05852381	15.87434615	14.00057761	8.37924654	8.808661536	14.06444721	14.61246434	2.18786217	1.714863584
TCEA1_1	4.072713367	5.239830254	17.02231468	15.07372937	8.665884628	8.986320437	14.37663942	15.67711761	2.137409715	2.048709104
TCEA1_2	3.736592278	4.930363627	15.13454889	14.25936174	8.452795197	8.90338965	14.39422983	15.94007919	2.419602884	2.132354932
TCEA1_3	4.208876868	4.784036531	15.44736408	14.37851558	8.652035628	8.672101949	13.64830628	14.80025312	1.991879162	2.103710604
TDG_1	4.382212092	4.969345557	15.49141767	14.44693439	8.696201968	8.736382163	13.5793023	15.00938464	2.327011395	2.281468885
TDG_2	4.406485142	4.706734251	15.48330184	14.44558112	8.450485184	8.748740442	14.22428852	15.51806471	2.118166717	1.986884941
TDG_3	4.807551586	5.296885952	16.36230149	14.69917229	9.064965656	9.169591661	14.49052319	15.54419231	2.044855874	2.203629781
TDP1_1	3.956783057	4.471672662	16.48288333	15.50733886	7.168522267	7.239765036	14.54575116	15.63345911	2.243635161	2.187932606
TDP1_2	4.518725629	5.142160083	16.30764117	15.07157524	9.46166456	9.354147536	16.37074056	18.03674216	2.690846094	2.608150616
TDP1_3	3.624509122	4.892966361	14.9270399	13.94349391	8.257187287	8.126424743	14.1011416	15.59470536	2.379560536	2.020040223
TDP2_1	4.476940382	5.036204344	16.29401533	14.77851063	8.136341949	8.289247482	14.29085651	15.60066481	2.219577575	2.211365627
TDP2_2	4.008506864	4.563335998	15.25870085	13.50987091	8.321091185	8.131452478	13.92334459	14.68299824	2.173789975	2.261933667
TDP2_3	4.503837719	5.050740071	16.02505171	13.73181767	8.580916414	8.69037177	14.50106336	15.51230445	2.333684466	2.203570477
TELO2_1	3.797012777	4.427146715	17.07067394	14.7842296	8.611663873	8.133315278	14.68861278	16.56970458	2.469886385	2.35816944
TELO2_2	4.506797217	5.008546574	16.87335779	14.57820946	8.87073231	8.643901016	14.35212415	15.19434223	2.734050228	2.071282749
TELO2_3	3.994160584	4.419117652	16.84759472	14.19461856	8.112810027	8.220033286	14.88764934	16.02267555	2.345914907	2.287442696
TERF1_1	4.473110774	4.845000762	15.60707306	14.35310821	8.746624911	8.832678131	14.16435984	15.01072348	2.100454103	2.000717004
TERF1_2	4.364863609	4.832756091	16.36275108	14.16488401	8.513215819	9.015858486	14.57604564	15.76318656	2.142249062	2.234223422
TERF1_3	4.364671573	5.232952121	16.51186937	14.49122448	8.786853635	8.544085119	13.75125292	15.33823276	2.056509043	2.189730036
TERF2_1	4.477906272	5.00354045	15.65716518	14.46466774	8.312238191	8.785899249	14.06007162	15.60336634	2.442026694	2.232547736
TERF2_2	4.396323646	5.162386862	16.73172513	15.02573007	8.594920287	8.992813192	14.35566198	15.33374614	2.310926894	2.240733504
TERF2_3	4.506292455	7.28890069	16.86142934	15.35143365	8.167031754	8.82231405	13.39464486	15.6180667 1	2.051781562	2.663226068
TFPT_1	4.213395452	4.948285124	15.45022687	13.84050629	8.277354927	8.597920474	13.86539768	15.37827734	2.029215674	2.247679162
TFPT_2	4.98043885	5.009107466	15.84986162	14.01479351	7.395901207	8.163877774	13.81626883	15.24857143	1.985664589	2.238161076

TFPT_3	4.117366619	4.7264532303	16.34527421	14.306664649	7.5220865594	7.5815126616	13.173762381	14.416586121	1.964972647	1.8603855431
THOC1_1	4.25510512	5.0056307191	15.364263281	12.868391981	9.3633830021	9.3285602221	15.563212871	17.038769151	2.740865555	2.8910322291
THOC1_2	4.658211693	5.5392759241	15.064997211	11.908575391	9.4246959781	9.446420741	15.167070771	16.457018431	3.441609971	2.4963593211
THOC1_3	4.537223341	5.4110986771	14.63332661	13.37886151	9.1191288341	9.0664272891	16.273387221	17.107443991	3.123656491	2.7044473131
THOC2_1	4.584806291	5.8642057791	15.043687241	15.037247761	8.7771962067	9.082771219	14.686558351	16.116110771	2.411108061	2.321290421
THOC2_2	4.525791261	5.6589750241	15.163270531	13.789614021	8.2671255451	8.9509189981	14.045257721	16.244306471	2.085655263	2.4361649541
THOC2_3	4.432565789	5.5467780161	15.104828221	13.856598271	8.4833669621	8.7767212791	14.544942171	15.823339041	2.320219052	2.0925819851
THOC3_1	4.201200343	4.829441161	14.974840831	14.346461651	8.543816731	8.8963618231	14.177521791	15.201507941	2.181500226	2.0939563171
THOC3_2	4.335883853	5.2669964097	15.688003761	15.534255151	8.4176354191	8.8246729091	14.035564921	15.899577791	2.310444747	1.6613350771
THOC3_3	4.57850680	5.178918785	16.358525591	15.294849281	9.4897311321	9.437417057	14.779421221	15.579570961	2.219675121	2.2501185861
THOC5_1	4.945214075	5.454977853	15.797428241	13.978913441	8.5416721711	9.0894659381	15.260668781	16.850418691	2.533314972	2.3855182091
THOC5_2	4.549783867	4.9531116791	16.732334221	13.569863871	9.1631772311	9.6604533321	16.786570741	18.530386761	3.297252368	3.2578355021
THOC5_3	4.620874219	3.9187227871	13.573274621	13.129671761	9.6044135751	10.027012741	17.119800371	18.133236981	3.909433392	3.9306246371
THOC6_1	4.28846581	4.7657024621	15.141266651	14.257788131	8.7597903921	9.1016072581	15.139077411	15.884931231	2.517881159	2.330909341
THOC6_2	4.223283793	4.6915457421	16.850361691	13.441544661	9.0261862991	9.2598863151	15.528604231	16.293586481	2.473224292	2.3325897231
THOC6_3	3.795503305	4.6941678621	15.812356661	13.699199241	7.9994650381	8.2630296981	14.237308191	15.314710451	2.156308476	1.9870114781
THOC7_1	3.330954081	5.4600676471	16.096770591	14.612934941	8.7798743471	9.9724386031	16.026200441	16.946301341	2.915882827	2.8231242791
THOC7_2	4.561250428	4.8466160911	15.527333721	13.655805651	9.2085533551	9.5346533381	15.284756761	16.695339661	2.859755733	2.1074684021
THOC7_3	4.055667138	5.3554605411	14.581316551	13.160410431	8.8987841271	9.9634721681	15.247619531	16.461895121	2.275187604	2.2254774431
THRAP3_1	4.578155657	5.751144941	19.026900281	15.698174631	9.5662385251	9.42161371	14.895427421	16.350485691	2.241971781	2.2923336631
THRAP3_2	5.69911316	5.9937882871	19.529896481	15.766816311	8.8656224871	9.4252081541	15.030455711	16.629113171	2.123633022	2.4253381551
THRAP3_3	5.41855351	6.0055726941	19.115738271	16.347934331	10.366306121	10.636029991	16.062056111	17.102111631	2.243920481	2.5713651011
TIMELESS_1	4.094172651	4.611774891	15.472580781	12.958393241	7.8395663431	8.5629751951	12.446212221	14.427149491	2.11546981	1.9388509971
TIMELESS_2	3.96900651	4.7858600681	15.509306821	13.469236251	7.5024563861	8.1151578781	12.274784591	14.230464891	2.301278488	1.9337738461
TIMELESS_3	4.06006161	4.4724804931	15.885402311	13.065572831	7.9996947311	8.5735572691	12.756508721	14.339720221	2.430502617	2.2136813251
TIPIN_1	3.225557726	4.7125750611	14.549997741	13.495485581	8.3032436881	8.8043551311	12.863419411	14.169576611	2.259335481	2.1732032811
TIPIN_2	3.079986905	3.8944587211	14.603441731	13.022067911	7.1557871211	7.4403289051	12.714186121	14.652597961	2.399728781	2.5000577111
TIPIN_3	3.950001539	4.4019118781	15.504781371	12.920441121	7.7951346891	7.3855882091	13.617123531	14.570759071	2.269721651	1.9435052311
TONSL_1	4.350855247	4.9709432281	14.560450951	14.066027091	8.1690441641	7.9994808821	12.689272771	15.310269631	2.429302282	1.3621642221
TONSL_2	5.203313661	5.6381190891	14.378258231	13.516151661	8.7385871061	9.1001883521	14.55734611	15.999792551	1.919817361	1.7998107041
TONSL_3	3.312191684	2.8557528581	15.408475321	13.60513471	8.5993418611	9.9226817361	11.973771741	13.630950381	2.67478064	1.8763134191
TOP1_1	4.346561254	4.6731428451	17.044127951	15.484657871	8.3209551941	8.7221232251	14.52111447	15.82908371	2.573812581	2.3419397551
TOP1_2	4.252953286	5.0023308821	16.914355351	15.985697021	9.1712061841	9.3494842531	14.550725641	16.270860421	2.211282382	2.1486273451
TOP1_3	4.314400812	4.9956387281	17.086400621	15.995243591	9.7082891861	10.005204621	15.798211461	16.896262531	2.340567441	2.2702156551
TOP2A_1	4.360894036	5.1267134491	14.914417791	13.768140571	9.735343231	9.8710669511	15.951779761	17.016216091	2.394557823	2.8424042981
TOP2A_2	4.810389073	5.1608253861	16.402436451	15.473741381	9.50657871	9.8373155641	15.578146991	16.826893561	2.642432861	2.5771785371
TOP2A_3	4.515009089	5.5170807121	15.792160051	13.640657661	9.4087880721	9.3367426871	15.73846551	17.036680371	2.571920196	2.2923145981
TOP2B_1	4.218928164	4.9183599431	16.177755291	13.853347271	8.0125095981	8.2234343711	14.455868741	15.967366171	2.258066709	2.1986016781
TOP2B_2	4.331114653	5.0981493691	15.865273441	14.696603431	8.5706552881	8.7239691681	14.208943331	15.722494581	2.215553104	2.029135691
TOP2B_3	4.553449681	5.1298222161	16.575301141	14.036667161	8.4866677721	8.8926058711	13.910361961	15.426298181	2.221388411	2.1991236121
TOP3A_1	4.594321625	4.7232866421	15.154675011	14.538834951	8.8666808491	8.9680343991	14.154656211	15.510022351	1.951090649	2.0246256751
TOP3A_2	4.450092491	4.8921193131	16.204841741	13.841692231	8.941659841	9.2729815481	14.436471951	15.548981781	2.340000753	2.2630255541
TOP3A_3	4.295504377	4.7505719621	16.046452491	14.205030371	8.8090332631	8.2547009711	15.068295051	15.554533641	2.240121144	2.0602823351
TOP3B_1	3.640303953	4.9388482361	15.734673071	13.742198361	7.9796896581	8.0580664721	14.344325151	15.676721041	2.151959821	1.9005771931
TOP3B_2	3.934053439	4.8344294451	16.469750411	13.63741339	8.0567462161	8.7160798391	13.711958261	15.178977571	2.082002038	2.0709943641
TOP3B_3	4.277653968	4.9792281291	16.096651191	15.053520621	7.1931917051	7.6831159591	12.936761881	13.619420441	2.000123841	1.9390830971

TOPBP1_1	4.525740514	5.357844607	15.74315808	13.74363553	8.660166692	9.263917021	14.47038208	15.59085201	2.275914672	2.055347287
TOPBP1_2	4.561788314	5.343601286	16.78978529	14.47092275	8.442830444	8.769426254	13.73146592	15.8522437	2.131651488	2.408502899
TOPBP1_3	4.492575704	4.539632309	16.87537371	14.11929089	7.059722549	7.07744874	13.56822833	14.85584414	2.656961526	2.358670968
TOPORS_1	4.319807714	5.001936624	15.58374762	14.16541875	8.491907905	8.566223385	14.41596406	16.02813863	2.166402266	2.024580219
TOPORS_2	4.666735346	5.08394521	17.47583898	14.6153277	8.032514729	8.711295799	13.58233068	14.72849032	2.356652302	2.314601809
TOPORS_3	4.627271184	5.469558915	16.82110271	14.33249824	8.560764561	8.548186197	14.20506171	15.77383096	2.136304612	1.79764256
TP53_1	4.234453087	5.004014622	15.78345285	13.94993668	8.548304029	8.548605509	14.09041623	15.15949641	2.454023546	2.003577544
TP53_2	4.079207332	4.846362836	17.35505246	13.88388995	8.648387019	8.56004766	13.93885366	15.10694257	2.161907576	1.930946935
TP53_3	4.140651113	4.742794008	15.27509976	14.7324902	7.709186839	7.758943161	13.90049554	14.65721182	2.085898457	2.188024876
TP53BP1_1_1	4.379988606	4.84578284	15.56991598	13.86899433	8.717214071	8.912151953	14.4774532	16.10337972	2.858973008	2.981466558
TP53BP1_1_2	3.903822775	4.543675642	15.52958145	14.14656947	7.996334549	7.957763789	14.6492466	15.50737471	3.533350502	3.360646037
TP53BP1_1_3	3.895695544	4.725746648	16.21246575	14.6445385	8.235613145	8.926381826	14.85194667	16.1423889	3.128362317	3.117194424
TP53BP1_2_1	4.524820226	5.183875099	16.44607	14.24059947	8.456289192	8.826619965	13.50910741	15.33160038	2.51329073	2.404695128
TP53BP1_2_2	4.111653183	4.750412699	16.11506842	15.0987915	8.862806138	9.544804695	14.72096318	16.10659762	2.458437591	2.155042707
TP53BP1_2_3	4.01798359	4.37758251	16.14826527	13.83574862	8.199637061	8.196721311	14.03940656	15.05526237	2.433433088	2.071911925
TP53BP2_1	4.163038047	5.005101907	15.82064298	14.37204349	8.339275454	8.724614662	14.08301788	15.41529868	2.206435544	2.258354191
TP53BP2_2	4.361782199	5.044562427	16.86526512	14.18329039	8.302889153	8.631142467	14.24287978	15.66121664	2.275405745	1.93697836
TP53BP2_3	4.439780026	5.060712637	16.20605091	14.13505893	8.039534365	8.406433626	14.10774917	14.51453861	2.063035282	2.029643237
TP53I11_1	2.856339612	3.334477691	8.285986214	7.229268293	9.265720481	9.969446361	17.05374378	18.41858096	3.481381101	4.224021167
TP53I11_2	4.119622485	4.604748967	15.72032768	13.68840911	7.96512022	8.777172249	14.1755937	15.92685231	2.146516101	1.979053707
TP53I11_3	3.95706916	4.638615136	16.1009378	13.87163082	8.045059995	9.287064363	13.92627442	15.10855234	2.22663638	2.359382977
TREX1_1	4.378606266	5.157809745	15.25490682	13.95158036	12.45521177	11.35103962	18.50738774	18.10017271	3.622857544	2.91817406
TREX1_2	4.387798234	5.007725112	16.43934685	15.100812481	12.56498746	11.38366303	19.552485	18.14943776	3.312479658	3.089868171
TREX1_3	4.364550968	5.223190906	14.60019165	14.42472308	11.77423945	10.90898534	18.57027983	18.06535516	3.747510385	2.954891661
TREX2_1	5.347655348	6.465184499	16.78366643	15.84756968	8.261074688	8.730265218	14.29232804	14.97975322	2.515594971	2.34375
TREX2_2	5.275687298	5.268873088	17.69625942	15.7511101	8.778108466	8.638782954	14.81027613	16.62213454	2.362723184	2.192816635
TREX2_3	5.471233727	5.403776617	17.24611399	16.7344232	8.19973578	8.844065167	14.99381643	16.49038498	2.778663508	2.380165289
TRRAP_1	3.672088264	4.663454065	16.10923416	14.51785103	9.011674069	9.630518964	14.96935862	16.31577103	2.324005904	2.260181388
TRRAP_2	3.649972026	4.361100032	15.81620949	13.92384854	8.629716591	9.005941034	14.52658605	16.15383944	2.205819851	2.255066674
TRRAP_3	3.641545455	4.283932038	16.01783763	14.30287805	8.707892873	9.075503831	14.21251544	15.43856467	2.234432117	1.927314962
UBA2_1	6.087629101	7.364972114	21.95690157	18.62253698	9.689981785	9.079565143	17.54925862	18.0116123	4.09228684	3.353156352
UBA2_2	6.222732491	7.572842999	22.11239329	17.41627473	10.05542396	10.16074642	16.2371401	17.93918461	2.804097525	2.724644167
UBA2_3	4.568488943	5.484496999	16.23833629	14.61441715	8.911482219	9.167552548	15.3663533	16.60586166	2.562110563	2.40525937
UBE2A_1	4.09235596	4.315662203	15.81447082	13.94293468	8.327383365	8.445578359	14.8857616	15.77839238	2.392919829	2.394620346
UBE2A_2	4.710967636	4.471779537	15.85380015	13.47736294	8.914040096	9.187537791	13.75911491	15.12485548	2.37473904	2.056857137
UBE2A_3	4.662237874	4.96370414	15.67873436	14.07101915	8.88280277	9.017284955	14.41676945	15.95143555	2.350800968	2.123369782
UBE2B_1	4.697790989	4.937796376	15.66900963	14.06292338	7.561795218	8.312221738	14.3975288	15.22355862	2.646172128	2.776454816
UBE2B_2	4.444568059	5.091962048	15.80859846	13.91375281	8.690899055	8.984736088	14.56190663	15.97835101	2.21563494	2.052540868
UBE2B_3	4.126195023	4.813127979	16.69154264	14.45873127	8.343260268	8.383646307	14.01582977	15.33166415	2.179670162	2.196493549
UBE2I_1	6.29694796	7.553214996	21.44927059	18.59748722	10.16520915	9.896396068	16.67049904	17.39806742	3.253676752	3.191334808
UBE2I_2	5.671673705	5.891526605	18.17574581	16.22117229	8.545908592	8.326660701	14.61276616	14.78730644	2.514958467	2.273666194
UBE2I_3	5.971229027	6.829229673	19.65109966	17.64868374	8.662279468	8.871446716	14.39117742	15.6993892	2.543450918	2.619056053
UBE2N_1	3.745637326	4.996891004	15.32048343	13.60725461	8.372408637	8.628927319	13.99854954	15.21178664	2.186909281	2.136306672
UBE2N_2	3.97809966	4.887700535	15.14599174	13.27880367	7.670087419	8.350547731	13.46946469	16.54860291	2.342844985	2.09908285
UBE2N_3	4.276398325	4.557027799	15.80677761	13.72535503	8.099000581	8.374960265	13.17275332	15.02268474	2.304760698	2.336727784
UBE2T_1	3.716031516	4.172575049	13.32517652	11.73244352	8.203189462	8.600717581	13.50611015	14.76681997	1.887012754	2.146851885

UBE2T_2	3.738449007	3.467387048	13.92056046	11.33008857	8.044639576	8.881143509	13.40759311	14.84421038	2.422924901	2.355290636			
UBE2T_3	3.63931966	3.7858007	13.62190018	11.57346679	8.00347579	8.308574407	13.84128613	14.32867741	2.545816403	2.2656			
UBE2V2_1	4.057205921	4.543509823	15.70988116	13.83718767	8.515508021	8.846691511	13.70721175	15.47308638	2.193135139	1.936364042			
UBE2V2_2	4.439472703	5.008405191	15.60709876	13.59026251	8.078842195	8.612047486	14.27717192	15.7146263	2.444493295	2.106524196			
UBE2V2_3	4.311627666	4.873735465	14.41377927	12.72864065	8.605609894	8.730135531	14.69780426	15.19161757	2.178337117	2.076617065			
UBN1_1	4.152819949	5.102784441	16.892909	14.92402975	7.955041674	8.419866679	13.70527607	15.19802822	1.95902696	2.289169521			
UBN1_2	4.213604116	4.814383909	16.12910896	13.82834674	8.334215563	8.746947211	13.78681661	15.00399317	2.249650742	2.030069365			
UBN1_3	4.418234578	4.891513125	16.54219589	14.31037563	9.242910325	9.497167423	15.35758676	16.54633371	2.937335516	2.416627553			
UCHL5_1	4.974286926	5.425194072	18.24207321	16.29551237	7.256366073	7.427759554	12.43829613	13.23421131	1.810013992	1.581494017			
UCHL5_2	4.97721226	5.597677177	18.3227975	16.30585929	7.462047066	7.787613988	12.45091926	13.03132656	1.956216908	1.803575163			
UCHL5_3	4.742837694	5.521062609	17.91687426	16.3334618	7.745218057	7.926510485	13.10547435	13.98674007	1.988008655	1.710255479			
UHRF1_1	4.035718715	4.708566567	15.73496811	13.8364665	8.321928831	8.296612576	13.6818899	15.43328602	2.134127282	2.341447907			
UHRF1_2	4.586880929	5.352599298	17.08486182	14.7464335	8.803640684	8.745024485	14.18230263	15.49153557	2.217996572	2.200768492			
UHRF1_3	4.269321633	4.832161455	13.53282266	13.37666926	8.630199743	8.872503129	14.64199933	16.43898564	2.004392291	2.108965262			
UIMC1_1	4.074773336	4.747778308	15.5842922	14.51266344	8.535393907	8.372312203	14.18195167	15.72265517	1.946220625	1.913445899			
UIMC1_2	4.266221369	5.302205672	16.20218252	14.5557656	8.562116445	8.562467132	14.74850174	15.71749654	2.472223592	2.36262422			
UIMC1_3	4.549809075	5.030687277	16.83668038	14.24071064	8.770773419	8.609784033	14.12996474	15.39862577	2.307407756	2.219429996			
UNG_1	4.090520346	4.74062353	16.00870181	13.35159555	8.379944429	8.164560506	13.91609575	14.7674619	1.955673265	2.231663783			
UNG_2	4.767846657	5.106393204	15.53824594	14.8719597	8.595452928	8.732902226	14.45702226	14.8779697	2.100217663	2.046103865			
UNG_3	4.424144049	4.296777713	15.93932759	14.73900596	8.047427151	7.81759924	13.74783255	15.10084133	2.117467076	2.049891357			
UPF1_1	3.475705468	4.331250187	14.31613936	12.07940212	8.438054852	8.529701164	13.7526987	15.02624204	2.103252962	2.060490176			
UPF1_2	3.865222319	4.068109198	14.86954956	13.26106629	8.271749455	8.696459927	14.31441896	15.80426116	2.235925647	2.199127772			
UPF1_3	4.339735178	4.847221606	16.2135739	13.68143945	7.941787942	8.686293824	14.69089139	15.72379274	2.519286832	1.974877718			
USP1_1	4.699391037	5.195280797	16.86473471	14.5886035	8.354999362	9.181570445	14.76500472	15.3808994	2.306955607	2.266995341			
USP1_2	4.307942904	5.260885109	16.53639489	15.34722551	8.599785481	8.343437065	13.43630873	16.07630225	2.060003123	2.128693274			
USP1_3	4.652749775	5.970628816	16.65731561	14.80301876	8.439961065	8.410303368	14.48972035	15.84383518	2.41885524	2.441606097			
USP28_1	4.214850203	4.804164506	16.47946855	13.91328395	7.406611138	7.143467591	14.45125028	15.36805628	2.210964871	2.076401466			
USP28_2	4.460265092	4.690886195	16.39858655	14.31350129	8.5540949	8.504245591	13.7735232	15.11543561	2.236682444	2.231094655			
USP28_3	4.44301792	5.262934474	16.73302164	14.06604467	8.216704668	8.707744535	14.07941365	15.39127852	2.361116595	1.951046889			
USP7_1	4.261422692	4.879432624	15.54728625	14.6477841	8.323465404	8.413854352	14.57242873	16.22005844	2.116141058	1.722917158			
USP7_2	3.896435285	4.511454599	15.21869078	13.46523901	8.023809364	7.979053804	13.13408646	14.86528025	1.997864078	1.967538373			
USP7_3	3.841648021	4.375678933	15.61484036	12.65660468	8.373874241	7.779938872	13.75186195	15.20694286	1.836042962	1.938730398			
UVSSA_1	3.984510792	4.666991413	15.77495562	12.76647514	8.659626649	8.554660105	14.06604914	15.33326907	2.008752159	2.147469735			
UVSSA_2	4.237182055	4.668394524	16.88874641	14.15570969	7.82850162	8.530961454	14.87070724	15.93940074	2.449360605	2.392220661			
UVSSA_3	4.41619729	5.293334215	16.62932362	14.71378887	8.263359493	8.945873097	14.30916323	16.09411215	2.422882782	2.444655009			
VPS72_1	6.80657245	7.857051055	19.69847647	17.36425774	8.894093902	8.907741251	14.03488548	15.99186033	2.512526957	2.704715564			
VPS72_2	6.630272953	7.577250855	17.77938692	17.71710991	9.257328449	9.106383095	14.52986546	16.67446642	2.66836445	3.041916995			
VPS72_3	4.621149877	5.410468455	17.48014031	15.49742816	7.937659323	7.936462238	13.72444022	15.03854252	2.350822746	2.263451554			
WAPL_1	5.575343969	6.26459298	20.84276115	19.15501091	8.167436287	8.626349603	14.18872566	15.30872619	2.246682303	2.335589111			
WAPL_2	5.466500344	6.120036538	21.14176863	18.68371761	8.403874	9.197635045	14.00768037	15.453621	1.899036819	2.07108433			
WAPL_3	4.632684507	5.514050502	19.04219279	15.97456787	8.633719158	8.690338199	14.36047981	15.46306998	2.238090484	2.13491142			
WEE1_1	4.669770148	6.130048697	19.55810147	13.71807731	8.802610896	8.520595117	15.05260224	16.11740614	2.079963023	2.45209149			
WEE1_2	4.260328627	5.584210164	17.82641601	16.84260151	8.532214397	8.544391232	14.66655563	15.58891536	2.056872721	2.114776205			
WEE1_3	4.559405022	5.422100206	16.62216882	14.48472971	11.38632963	11.50336272	16.8532628	20.62477689	3.081998886	2.659458576			
WRN_1	4.061061268	4.76899092	16.57999531	14.62712363	7.899345608	7.941586674	13.93212777	15.31074032	2.10647795	1.968955854			
WRN_2	4.310423245	4.678265933	16.57691635	14.44187009	7.244751384	7.495393249	13.01245626	14.21623232	2.149667695	2.157484723			

WRN_3	4.008118518	4.849091546	15.19665872	14.53705515	8.187849447	8.482375761	13.81010595	15.47042833	2.121125261	2.065720081
XAB2_1	4.950709566	5.265765447	14.74375821	16.0519845	8.273254051	8.432329111	13.96056828	14.31461618	2.050483421	2.206292702
XAB2_2	4.33290563	5.074370709	17.05595925	14.75462825	8.495539297	9.068219005	13.90852376	15.35689821	2.223604097	2.044985423
XAB2_3	3.656700366	4.029605263	15.09681323	14.62293453	7.734578051	8.433768853	13.40871711	14.94581003	2.208364777	2.247646362
XPA_1	3.98848154	4.746872154	15.87901027	14.13612565	8.652365544	9.168280588	14.40192157	16.1059816	2.416498505	2.106045591
XPA_2	3.837075135	4.579161614	16.59340999	14.92138316	7.401238096	7.993991989	14.30651341	14.99714657	1.977229516	2.383993569
XPA_3	3.640540779	4.216495919	15.48142376	14.48437769	7.07361207	7.195657245	13.69579225	15.03607945	2.038282462	2.113149314
XPC_1_1	3.966648083	4.674555179	16.73532205	15.06006721	8.415029177	8.405089647	14.37404843	15.93469992	2.221368822	2.182754981
XPC_1_2	4.404536862	4.53865697	16.88080667	14.8177596	6.915052992	7.663822251	14.31629852	15.29309861	2.053012174	2.464225757
XPC_1_3	3.438726398	3.767341466	15.49602073	13.7775204	7.658377964	8.346285655	14.0827891	15.50795785	2.332707473	2.213360148
XPC_2_1	4.499564982	5.069903975	15.66524126	14.61222016	8.792921823	8.314881271	13.67387592	14.9121356	2.10798709	2.131814483
XPC_2_2	4.467788264	4.661394174	15.16691742	13.79684288	8.524637035	8.97089709	14.30783314	15.54260911	2.284677261	2.129434354
XPC_2_3	4.313867696	4.841627415	15.29739216	14.56266462	8.462295333	8.534761745	14.34374672	15.26559941	2.27831503	2.434648524
XRCC1_1	4.678431488	5.385423639	15.56115183	14.15419829	8.716052755	8.748156969	14.00811641	14.72065012	2.533030236	2.525052369
XRCC1_2	4.179403954	4.565136694	17.18618202	15.17926913	7.776514349	7.986868228	13.54338182	14.16449656	2.278949053	2.263741375
XRCC1_3	4.200681396	5.077185125	16.02929631	13.98226742	8.340373431	8.781251346	13.90279757	15.59655907	2.07034215	2.096841531
XRCC2_1	4.599904881	5.057324462	16.02470993	15.02930627	9.066504663	9.183862712	14.22483961	15.37549046	2.030481105	2.555586383
XRCC2_2	4.410247307	4.653432466	16.12148961	14.60588951	8.302767721	9.044407396	14.22126338	15.16902932	2.326102326	2.356155614
XRCC2_3	3.962537736	4.38025167	15.7758297	13.54045455	7.699222456	7.944263302	13.27796452	15.48273289	2.236263936	2.312610686
XRCC3_1	3.103000977	4.186778985	15.50348731	12.7302256	7.325299213	7.949221702	13.08397653	14.81743742	2.413596788	2.149519869
XRCC3_2	4.116429679	4.744687404	15.88357082	14.46113117	8.069443172	7.857472607	13.70200398	15.56926747	2.336849543	2.204512874
XRCC3_3	4.550230047	5.427495292	18.00559883	15.43894609	7.491356355	8.217808029	13.26103416	14.98253354	2.116432074	2.248041546
XRCC4_1	4.07332692	4.68426548	15.35254498	14.043652	8.598608381	8.667494771	14.26872697	15.50947278	2.275932613	2.369141926
XRCC4_2	4.009298707	4.809976625	16.30496144	14.39224633	8.763604907	8.7485745 22	14.37001746	15.57267548	2.267172303	2.218388452
XRCC4_3	3.70412485	4.22034325	15.76041492	13.91270965	7.569875124	8.073093608	13.74282513	14.6867202	2.106046654	2.248777916
XRCC5_1	6.364185175	6.538101718	19.6365454	16.01093207	9.352770343	9.818670031	15.96966016	17.44440179	2.684571741	2.726331713
XRCC5_2	7.232127979	6.987601805	18.66689993	17.51432826	10.09671865	10.17817449	16.61605384	18.55565157	2.87134472	3.24985361
XRCC5_3	6.101458638	6.512834652	19.59715566	17.37271381	8.201372868	9.15906977	15.64846146	18.56304551	3.385245424	3.196654372
XRCC6_1	8.033465093	6.91107924	19.24086677	18.54807766	9.517893087	10.00138539	15.98251739	17.1603197	3.194073801	2.914654539
XRCC6_2	6.30558577	6.307076101	20.14975564	16.43649877	9.345621209	10.000756	16.07865963	17.73976472	2.742475453	2.728514736
XRCC6_3	5.34892316	6.504698847	18.8259304	16.54237093	8.74164637	8.520289322	14.12133397	16.39496047	2.112060302	1.964902345
YBX3_1	4.709826774	4.624013786	16.05518665	13.11726395	7.634533903	8.46662738	13.47912465	14.91391332	2.000286985	2.029973657
YBX3_2	4.011571904	5.037904617	17.36727489	14.84348488	8.482885305	9.110230498	14.50519415	15.98440222	2.316274309	2.136748653
YBX3_3	3.546673336	4.335422491	14.90461 65	13.33581008	8.808602986	8.743187412	15.03480606	16.57150805	2.633753512	2.044271002
YEATS4_1	4.289689612	5.037020116	16.81323465	16.08368975	8.439050343	8.463834912	14.08778343	15.5803584	2.254977859	2.426391721
YEATS4_2	4.124685139	5.080432638	15.54305633	13.43583769	9.990309754	9.933246825	16.01668497	18.2764669	3.561408928	3.430294595
YEATS4_3	4.617037085	5.009698492	17.16474968	15.67490042	8.817022284	8.543778136	14.53460689	15.3047415	2.664793372	2.215208034
YY1_1	4.183431627	4.783825361	16.28645256	14.1260449	8.205784309	8.348502276	14.16681642	15.2400535	2.215623091	2.169631685
YY1_2	4.114784974	4.732459064	15.42421077	14.3061971	8.295106327	8.220213978	13.97798126	15.70671086	2.024221692	1.830786613
YY1_3	4.385790832	4.975112984	15.81171014	14.60008937	8.615217607	8.853765099	13.75443972	15.35271219	2.137758427	2.18155115
ZNHIT1_1	7.739959563	8.240746127	22.32771138	19.98335731	8.354017496	8.205261112	14.13796963	15.32474208	2.222361651	2.33398546
ZNHIT1_2	4.100248779	5.038880916	14.94229092	14.5013816 9	8.611518738	8.827267736	14.41832976	15.31183035	2.305915378	2.293406604
ZNHIT1_3	4.115900345	5.134457502	16.66192236	14.16784377	8.76125222	9.139842479	14.19815971	15.48998742	2.001795044	2.386240862
ZRANB3_1	4.828722883	5.249112768	16.33396988	14.60499952	8.491176338	8.840569082	15.01465457	15.85280169	2.082828212	2.214396774
ZRANB3_2	4.662641439	4.818881146	17.06053337	15.75284477	7.866512017	8.245496411	14.28693279	15.58913696	2.095161811	2.098235575
ZRANB3_3	4.056099595	5.004928757	15.74573164	14.39500297	8.254106745	8.384818731	14.56996072	15.47310602	2.265396383	2.237410072

non-targeting_00026	4.420687166	5.346218636	16.41540604	14.92318009	8.802410521	9.844117019	14.29125697	16.2652761	2.371703128	2.331023639			
non-targeting_00074	4.272501453	4.76621888	15.36855057	14.76987499	9.035243228	9.439198882	16.21021367	17.7521414	2.767509306	2.456336064			
non-targeting_00097	4.162231009	4.786692782	15.91202324	13.96204092	7.883955265	8.710913304	13.34571006	15.17975529	1.959684832	2.021876922			
non-targeting_00155	4.199866928	5.066880778	15.86788646	13.6184242	8.266810694	9.164550457	14.3721063	15.33752087	1.973763982	1.901339231			
non-targeting_00240	4.80021574	5.130354551	16.19625906	13.6513303	8.784708894	9.108435081	14.68106423	15.7751946	2.271561124	2.167658602			
non-targeting_00249	4.361231644	4.963697455	16.32542111	13.64200623	8.681320064	8.682361529	13.96389544	15.0833252	2.307995163	1.887199396			
non-targeting_00364	4.289091059	5.102741214	15.79204024	14.62669028	8.392591448	8.675150113	14.11173504	15.32657292	2.266764162	1.961922605			
non-targeting_00398	4.356339537	5.124290181	16.4882336	14.36842649	8.27294686	8.206181666	13.96809747	15.29069626	2.499063467	2.039577138			
non-targeting_00513	4.422090116	4.993104133	15.8733169	14.30832151	8.512946676	8.943097024	14.37442245	15.21145624	2.244496329	2.010306359			
non-targeting_00515	4.123402675	4.876832255	14.38260556	13.58040236	8.586489606	8.881519687	14.33665458	15.09657424	2.099787567	2.123326506			
non-targeting_00566	4.342130564	4.861267934	15.84836279	14.77042317	8.405454402	8.590210206	14.35559526	16.12144047	2.122987877	2.180988161			
non-targeting_00678	4.100810431	5.076946114	15.45289634	14.84240643	8.209735734	8.549433607	13.77239057	16.00856769	2.224618033	2.023876683			
non-targeting_00740	3.943594397	4.335246841	16.45112351	14.12722881	7.408822786	7.807591643	14.15730909	15.44057948	2.32896633	2.166601332			
non-targeting_00745	4.806512366	5.169698434	16.80147892	14.06527686	8.936845602	9.369510268	14.12224677	16.08360852	2.674721173	2.274504117			
non-targeting_00869	4.427639635	5.291541002	16.50612144	14.29590353	8.749465999	9.044585887	14.24871566	15.37068861	2.196851648	2.234249086			
non-targeting_00948	4.102387836	4.452063303	14.91032183	14.37983762	7.553461389	8.287045185	13.39775998	15.05437211	2.125949737	2.012938891			
non-targeting_00984	4.866929134	5.067701569	15.40941285	13.70769343	8.876267507	8.975742687	14.43248608	15.68381977	2.206038657	2.275340103			
non-targeting_01042	3.861654576	4.960544546	17.70483174	14.06446387	8.508793257	9.089112042	14.12457381	16.19780299	1.976872083	2.018842019			
non-targeting_01091	4.53620458	4.804655347	15.92424767	14.5872845	8.404175538	8.432217435	14.39994325	15.77772167	2.201936782	1.928504233			
non-targeting_01222	4.065262063	5.226919026	16.6711783	14.14954076	8.768673403	8.588026704	14.56771792	15.74190127	2.319497468	2.238258377			
non-targeting_01332	4.319879006	5.058354174	15.92427617	14.05779691	8.616811575	9.142194127	14.40624635	16.04289444	2.306897265	2.098140285			
non-targeting_01358	4.329632052	5.068415303	16.22157062	15.00033546	8.740565015	8.513658015	13.97941681	15.46157817	2.181364911	2.320294457			
non-targeting_01439	4.472598138	4.72887768	17.34701259	15.15736786	8.721024085	7.912574353	14.66637011	15.91087116	2.455710788	1.741923809			
non-targeting_01528	4.755981368	4.430842925	16.86315874	15.16830522	8.895802052	8.764589298	14.48625947	15.61084216	2.291424661	1.734507313			
non-targeting_01566	3.977680351	4.372442019	15.60294014	12.97295438	8.263675763	8.704509511	14.47331427	15.02259294	2.288145663	2.009775646			
non-targeting_01594	4.481632653	4.975164004	15.85408151	14.51881628	8.802982201	8.98989899	14.26903844	15.11227911	2.45748742	2.170026581			
non-targeting_01631	4.342872924	4.826412681	16.59798535	14.41725672	7.949632102	8.568740033	13.67853146	15.34791677	2.149624644	1.783134851			
non-targeting_01959	4.672870476	5.024553731	15.84104637	15.11533483	8.843176944	8.753566982	14.43051573	15.53822318	2.289190576	2.303507498			
non-targeting_02031	4.850166216	5.119118523	14.36257073	14.63407852	8.869958	8.886696944	15.32385235	16.7994117	2.470662636	2.417596791			
non-targeting_02085	3.850911897	4.867349777	15.63380602	15.03077353	7.719595707	7.825973216	13.24581515	14.62018891	2.315006201	2.386795074			
non-targeting_02293	4.273774963	4.886248143	17.10125601	14.83931654	8.5656902	8.744220246	14.24047129	15.50741366	2.168921468	1.880779392			
non-targeting_02305	4.054347153	5.23170179	15.36621474	14.24479581	8.566787635	9.045232507	14.01876551	15.13509085	2.31587812	2.138187149			
non-targeting_02359	4.453352937	4.816224902	16.15193936	13.43731776	8.577282907	8.868996847	14.44587956	15.66657363	2.144152826	1.971154376			
non-targeting_02607	4.218836014	4.81291111	16.02599792	14.02691092	8.473858816	8.734291534	14.18089769	15.36581631	2.230586225	2.188380791			
non-targeting_02613	4.01840457	4.951790399	15.77972249	13.21785509	8.425917065	9.17371674	13.94027137	15.28548416	2.094374992	2.208658186			
non-targeting_02642	3.766559755	5.206598115	16.05489062	13.38660834	8.957569918	8.800633173	13.76483441	15.2964127	2.399085025	2.246369393			
non-targeting_02692	4.733631251	5.040133399	16.16090859	13.97005182	9.023732286	9.38256757	14.60153167	15.78546668	2.142263307	2.386301477			
non-targeting_02821	4.83195947	5.979340066	15.98294647	14.44094887	9.05618502	8.92613934	14.51166793	17.24621514	2.315228801	1.836923077			
non-targeting_02822	4.466696359	5.072149327	16.51452282	14.97979549	8.266858598	8.704306255	14.46246424	15.63885171	2.498660725	2.166053763			
non-targeting_02937	4.66023512	5.196957174	16.13644956	14.7708831 5	9.392123477	9.61942831	14.9901901	16.37516021	2.357586803	1.872504799			
non-targeting_02942	4.363367223	5.04656813 1	15.55660172	15.13275466	8.439047554	8.734673376	14.13491657	16.16753898	2.20772222	2.247120445			
non-targeting_02964	4.1742 23	4.668717073	16.12595566	14.22370469	8.631919683	8.761152235	14.33295128	16.12744272	2.437577745	2.389737466			
non-targeting_03019	4.220096608	5.423060657	17.43682899	14.51212132	9.216793338	9.296026194	14.90163311	15.93112514	2.134637473	2.197728567			
non-targeting_03026	3.956687786	4.968287585	16.07041102	14.11784706	7.549956112	8.260775566	14.26729386	14.9170726	2.237042797	2.306861917			
non-targeting_03051	3.928475658	4.387217343	14.61905053	13.31653356	6.944156929	7.194990056	12.04455662	13.05464713	1.848696128	1.682274944			

non-targeting_03659	4.193624082	4.951513219	16.76617761	14.88626095	8.556591971	8.537666174	13.68742228	16.07448987	2.176846101	2.136252228
non-targeting_03676	3.974018935	4.454865067	14.68429714	14.17393099	8.439786431	8.784360678	13.78159658	15.05392758	2.158812099	2.03621144
non-targeting_03686	4.210291548	4.602637371	16.62691935	14.42531443	8.206620067	9.011643085	14.05079152	15.60285499	2.364094953	2.192443653
non-targeting_03178	4.353432636	5.219833431	16.71983661	14.47153803	8.314669403	8.110026619	14.47286838	15.40781185	1.954446205	2.086145405
non-targeting_03220	4.265033266	5.055760211	17.32421437	14.7525608	8.559094211	8.795339899	14.56627761	16.99502372	2.3447691	2.233257354
non-targeting_03260	4.491886995	5.256684196	16.10905603	14.79637659	8.862725206	9.047083234	14.50278787	15.60886964	2.543376607	1.976185424
non-targeting_03264	4.317996655	4.850632066	16.33452666	14.16189483	8.577063387	9.219537715	14.01557662	15.92664307	2.048376362	1.935388023
non-targeting_03292	4.230081189	4.674605637	14.87481184	14.044794291	8.618861974	8.447590555	13.97145874	14.12945818	2.405646786	2.256006586
non-targeting_03346	4.584448634	5.540393397	17.18994666	14.41555696	8.920262451	9.371973129	15.32820179	16.2613806	2.215124685	1.819696275
non-targeting_03451	3.602764368	4.508813206	16.56181034	13.80454073	8.022215791	8.136437541	14.56891961	15.84950817	2.020605184	2.114395239
non-targeting_03460	4.056662776	4.528196315	15.48808752	13.70017868	8.254629931	8.252840198	13.76246381	16.51612503	2.462683612	3.140527456
non-targeting_03512	3.660173514	4.355350737	15.78822317	14.02650803	7.043382786	7.136421732	13.16210942	14.90478541	2.136791496	2.132974782
non-targeting_03589	4.189476542	4.904587351	16.30274253	14.83861029	8.709357078	9.044736031	14.1262442	15.65246675	2.279918945	1.979319152
non-targeting_03612	4.426771024	5.296166318	16.48177496	13.91744367	8.974340789	8.650786675	14.30840909	15.21612339	2.211452256	2.20400767
all_non_targeting	4.267461342	4.931927742	16.06786202	14.29308141	8.458498055	8.714047665	14.20807201	15.57740672	2.249155446	2.125341641

Table 7

Figure	Cell type	Fastq file name	Prime editor architecture	Condition	Sequenced amplicon(s)	pegRNA target site	Intended edit	Complementary strand nick position	pegRNA spacer	pegRNA scaffold (5'-3')	pegRNA RT template (flap sense, 5'-3')	pegRNA 3' extension (5'-3')	pegRNA linker	pegRNA 3' structural motif (5'-3')	sgRNA spacer (5'-3')	sgRNA scaffold (5'-3')
40C,40D,40E,40G,41E,47B,47D,47E,47F	HeLa integrase library	Chen2021_K562_PE2_screen_pPC1000	SaPE2	Repeat-seq of K562 PE2 screen	pPC1000 sgRNA and edit site	pPC1000 site	+6 G to C		GTCTGATAGGCAGCCTGCACGTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCACGTG (SEQ ID NO: 5184)	TGGTGCGGTGAATTC (SEQ ID NO: 5185)	GAATTCACCGCACCAGTGCAGGCGT (SEQ ID NO: 5186)	0	0	0	0
40D,40F,40G,41F,41G,41H,47C,47G,47H	HeLa integrase library	Chen2021_K562_PE2_sem3+50_integration_pPC1000	SaPE2	Repeat-seq of K562 PE3+50 screen	pPC1000 sgRNA and edit site	pPC1000 site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACGTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCACGTG (SEQ ID NO: 5184)	TGGTGCGGTGAATTC (SEQ ID NO: 5185)	GAATTCACCGCACCAGTGCAGGCGT (SEQ ID NO: 5186)	0	0	GGTTCAACAAAATATTGTCGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)
40D,40E,47E,47F	HeLa CRISPR/cas9 cells with integrase library	Chen2021_HeLa_PE2_screen_pPC1000	SaPE2	Repeat-seq of HeLa PE2 screen	pPC1000 sgRNA and edit site	pPC1000 site	+6 G to C		GTCTGATAGGCAGCCTGCACGTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCACGTG (SEQ ID NO: 5184)	TGGTGCGGTGAATTC (SEQ ID NO: 5185)	GAATTCACCGCACCAGTGCAGGCGT (SEQ ID NO: 5186)	0	0	0	0
40D,40F,41G,41H,47H,47I	HeLa CRISPR/cas9 cells with integrase library	Chen2021_HeLa_PE2_sem3+50_integration_pPC1000	SaPE2	Repeat-seq of HeLa PE3+50 screen	pPC1000 sgRNA and edit site	pPC1000 site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACGTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCACGTG (SEQ ID NO: 5184)	TGGTGCGGTGAATTC (SEQ ID NO: 5185)	GAATTCACCGCACCAGTGCAGGCGT (SEQ ID NO: 5186)	0	0	GGTTCAACAAAATATTGTCGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)
40D,47A,47I	HeLa CRISPR/cas9 cells with integrase library	Chen2021_Hela_PE3-50_screen_pPC1000	SaPE2	Repeat-seq of HeLa PE3-50 screen	pPC1000 sgRNA and edit site	pPC1000 site	+6 G to C	-50	GTCTGATAGGCAGCCTGCACGTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCACGTG (SEQ ID NO: 5184)	TGGTGCGGTGAATTC (SEQ ID NO: 5185)	GAATTCACCGCACCAGTGCAGGCGT (SEQ ID NO: 5186)	0	0	GAGAACGCTCCTTCAACAAATCTGC (SEQ ID NO: 5188)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)
46A	HEK293T with integrase library	Chen2021_0001	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACGTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCACGTG (SEQ ID NO: 5184)	TGGTGCGGTGAATTC (SEQ ID NO: 5189)	AATTCACCGCACCAGTGCAGGCGT (SEQ ID NO: 5190)	0	0	0	0
46A	HEK293T with integrase library	Chen2021_0002	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACGTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	CAGCCTGCACGTG (SEQ ID NO: 5191)	TGGTGCGGTGAATTC (SEQ ID NO: 5189)	AATTCACCGCACCAGTGCAGGCGTG (SEQ ID NO: 5192)	0	0	0	0
46A	HEK293T with integrase library	Chen2021_0003	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACGTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCAGCCTGCAC (SEQ ID NO: 5193)	TGGTGCGGTGAATTC (SEQ ID NO: 5189)	AATTCACCGCACCAGTGCAGGCGTGC (SEQ ID NO: 5194)	0	0	0	0

46A	HEK293T with integrated HBB CDS	Chen2021_0004	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GGCAGGCTGCACTGGTGGGTGAATT (SEQ ID NO: 5198)	AATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5135)	0	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0005	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGGCAGGCTGCACTGGTGGGTGAATT (SEQ ID NO: 5197)	AATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5199)	0	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0006	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	TAGGCAGGCTGCACTGGTGGGTGAATT (SEQ ID NO: 5188)	AATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5200)	0	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0007	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGGCCTGCACTGGTGGGTGAATT (SEQ ID NO: 5184)	AGAATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5202)	0	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0009	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	CAGGCCTGCAC (SEQ ID NO: 5191)	AGAATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5203)	0	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0009	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCAGGCCTGCAC (SEQ ID NO: 5193)	AGAATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5204)	0	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0010	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GGCAGGCCTGCAC (SEQ ID NO: 5196)	AGAATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5205)	0	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0011	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGGCAGGCTGCACTGGTGGGTGAATT (SEQ ID NO: 5197)	AGAATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5206)	0	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0012	SaPE2	PE2	integrated HBB CDS	integrated HBB CDS	+6 G to C	GTCTGATAGGCAGCCTGCATGAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	TAGGCAGGCTGCACTGGTGGGTGAATT (SEQ ID NO: 5189)	AGAATTCAGCGGCAGCAGTGCAGGCTGCCTCA (SEQ ID NO: 5207)	0	0	0	0	0

46A	HEK293T with integrated HBB CDS	Chen2021_0013	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATTCTTT (SEQ ID NO: 5208)	AAAGAATTCACCGCACCAGTGCAGGCT (SEQ ID NO: 5209)	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0014	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	CAGCCTGCAC (SEQ ID NO: 5191)	TGGTGCGGTGAATTCTTT (SEQ ID NO: 5208)	AAAGAATTCACCGCACCAGTGCAGGCTG (SEQ ID NO: 5210)	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0015	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCAGCCTGCAC (SEQ ID NO: 5193)	TGGTGCGGTGAATTCTTT (SEQ ID NO: 5208)	AAAGAATTCACCGCACCAGTGCAGGCTGC (SEQ ID NO: 5211)	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0016	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GGCAGCCTGCAC (SEQ ID NO: 5195)	TGGTGCGGTGAATTCTTT (SEQ ID NO: 5208)	AAAGAATTCACCGCACCAGTGCAGGCTGCC (SEQ ID NO: 5212)	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0017	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGGCAGCCTGCAC (SEQ ID NO: 5197)	TGGTGCGGTGAATTCTTT (SEQ ID NO: 5208)	AAAGAATTCACCGCACCAGTGCAGGCTGCCT (SEQ ID NO: 5213)	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0018	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	TAGGCAGCCTGCAC (SEQ ID NO: 5199)	TGGTGCGGTGAATTCTTT (SEQ ID NO: 5208)	AAAGAATTCACCGCACCAGTGCAGGCTGCCTA (SEQ ID NO: 5214)	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0019	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATTCTTTGC (SEQ ID NO: 5215)	GCAAAGAATTCACCGCACCAGTGCAGGCT (SEQ ID NO: 5216)	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0020	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	CAGCCTGCAC (SEQ ID NO: 5191)	TGGTGCGGTGAATTCTTTGC (SEQ ID NO: 5215)	GCAAAGAATTCACCGCACCAGTGCAGGCTG (SEQ ID NO: 5217)	0	0	0	0
46A	HEK293T with integrated HBB CDS	Chen2021_0021	SaPE2	PE2	Integrated HBB CDS	Integrated HBB CDS	+6 G to C		GTCTGATAGGCAGCCTGCACTGG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCAGCCTGCAC (SEQ ID NO: 5193)	TGGTGCGGTGAATTCTTTGC (SEQ ID NO: 5215)	GCAAAGAATTCACCGCACCAGTGCAGGCTGC (SEQ ID NO: 5218)	0	0	0	0

46A	HEK293T with integrase at HBB CDS	Chen2021 0022	SaPE 2	PE2	integrated HBB CDS	integrated HBB CDS	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GGCAGGCCTGCAC (SEQ ID NO: 5196)	TGGTGCGGTGAATT CTTTGC (SEQ ID NO: 5215)	GCAAAGAATTCACCGCA CCAGTGCAGGCTGCC (SEQ ID NO: 5219)	0	0	0	0
46A	HEK293T with integrase at HBB CDS	Chen2021 0023	SaPE 2	PE2	integrated HBB CDS	integrated HBB CDS	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGGCAGGCCTGCAC (SEQ ID NO: 5197)	TGGTGCGGTGAATT CTTTGC (SEQ ID NO: 5215)	GCAAAGAATTCACCGCA CCAGTGCAGGCTGCCT (SEQ ID NO: 5220)	0	0	0	0
46A	HEK293T with integrase at HBB CDS	Chen2021 0024	SaPE 2	PE2	integrated HBB CDS	integrated HBB CDS	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	TAGGCAGGCCTGCAC (SEQ ID NO: 5198)	TGGTGCGGTGAATT CTTTGC (SEQ ID NO: 5215)	GCAAAGAATTCACCGCA CCAGTGCAGGCTGCCTA (SEQ ID NO: 5221)	0	0	0	0
46C	Hela CRISPRi cells with integrase at pPC1000 (non-H6	Chen2021 0025	SaPE 2	PE2	pPC1000 edit site	pPC1000 edit site	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GGCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAAT (SEQ ID NO: 5223)	ATTCACCGCACCAGTGC AGGCC (SEQ ID NO: 5224)	0	0	0	0
46C	Hela CRISPRi cells with integrase at pPC1000 (non-H6	Chen2021 0026	SaPE 2	PE2	pPC1000 edit site	pPC1000 C edit site	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAAT (SEQ ID NO: 5223)	ATTCACCGCACCAGTGC AGGCT (SEQ ID NO: 5225)	0	0	0	0
46C	Hela CRISPRi cells with integrase at pPC1000 (non-H6	Chen2021 0027	SaPE 2	PE2	pPC1000 edit site	pPC1000 C edit site	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT (SEQ ID NO: 5189)	AATTCACCGCACCAGTG CAGGC (SEQ ID NO: 5226)	0	0	0	0
46C	Hela CRISPRi cells with integrase at pPC1000 (non-H6	Chen2021 0028	SaPE 2	PE2	pPC1000 edit site	pPC1000 C edit site	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT (SEQ ID NO: 5189)	AATTCACCGCACCAGTG CAGGCT (SEQ ID NO: 5130)	0	0	0	0
46C	Hela CRISPRi cells with integrase at pPC1000 (non-H6	Chen2021 0029	SaPE 2	PE2	pPC1000 edit site	pPC1000 C edit site	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT (SEQ ID NO: 5189)	AATTCACCGCACCAGTG CAGGCT (SEQ ID NO: 5227)	0	0	0	0
46C	Hela CRISPRi cells with integrase at pPC1000 (non-H6	Chen2021 0030	SaPE 2	PE2	pPC1000 edit site	pPC1000 C edit site	H6 G to C	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT C (SEQ ID NO: 5195)	GAATTCACCGCACCAGT GCAGGCT (SEQ ID NO: 5195)	0	0	0	0

46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0031	SaPE 2	PE2	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACAGTG GCAGGC (SEQ ID NO: 5228)	0	0	0	0
46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0032	SaPE 2	PE2	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACAGTG GCAGGCT (SEQ ID NO: 5202)	0	0	0	0
46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0033	SaPE 2	PE3+50	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAAT (SEQ ID NO: 5223)	AATTCACCGCACAGTGC AGGC (SEQ ID NO: 5224)	10	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0034	SaPE 2	PE3+50	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAAT (SEQ ID NO: 5223)	AATTCACCGCACAGTGC AGGCT (SEQ ID NO: 5225)	10	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0035	SaPE 2	PE3+50	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT (SEQ ID NO: 5189)	AATTCACCGCACCAGTG CAGGC (SEQ ID NO: 5226)	10	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0036	SaPE 2	PE3+50	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT (SEQ ID NO: 5189)	AATTCACCGCACCAGTG CAGGCT (SEQ ID NO: 5190)	10	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0037	SaPE 2	PE3+50	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT C (SEQ ID NO: 5195)	GAATTCACCGCACCAGT GCAGGC (SEQ ID NO: 5227)	10	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0038	SaPE 2	PE3+50	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT C (SEQ ID NO: 5195)	GAATTCACCGCACCAGT GCAGGCT (SEQ ID NO: 5186)	10	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPR cells with integrated pPC1000 insert	Chen2021 0039	SaPE 2	PE3+50	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	TCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACCAGTG GCAGGC (SEQ ID NO: 5228)	10	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)

46C	HeLa CRISPRi cells with integrated pPC1000 linch-1	Chen2021_0040	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGCT (SEQ ID NO: 5202)	10		GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPRi cells with integrated pPC1000 linch-2	Chen2021_0041	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	ATTCACCGCACGAGTGC AGGC (SEQ ID NO: 5224)	10		GAGAACCGTC CTAACAAAT CTGC (SEQ ID NO: 5186)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPRi cells with integrated pPC1000 linch-3	Chen2021_0042	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AATTCACCGCACGAGTGC AGGCT (SEQ ID NO: 5223)	10		GAGAACCGTC CTAACAAAT CTGC (SEQ ID NO: 5186)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPRi cells with integrated pPC1000 linch-4	Chen2021_0043	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AATTCACCGCACGAGTG CAGGC (SEQ ID NO: 5225)	10		GAGAACCGTC CTAACAAAT CTGC (SEQ ID NO: 5186)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPRi cells with integrated pPC1000 linch-5	Chen2021_0044	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CAGGCT (SEQ ID NO: 5180)	AATTCACCGCACGAGTG CAGGCT (SEQ ID NO: 5180)	10		GAGAACCGTC CTAACAAAT CTGC (SEQ ID NO: 5186)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPRi cells with integrated pPC1000 linch-6	Chen2021_0045	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT C (SEQ ID NO: 5195)	GAATTCACCGCACGAGT GCAGGC (SEQ ID NO: 5227)	10		GAGAACCGTC CTAACAAAT CTGC (SEQ ID NO: 5186)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPRi cells with integrated pPC1000 linch-7	Chen2021_0046	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT C (SEQ ID NO: 5195)	GAATTCACCGCACGAGT GCAGGCT (SEQ ID NO: 5188)	10		GAGAACCGTC CTAACAAAT CTGC (SEQ ID NO: 5186)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPRi cells with integrated pPC1000 linch-8	Chen2021_0047	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	GCCTGCAC (SEQ ID NO: 5222)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGC (SEQ ID NO: 5226)	10		GAGAACCGTC CTAACAAAT CTGC (SEQ ID NO: 5186)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46C	HeLa CRISPRi cells with integrated pPC1000 linch-9	Chen2021_0048	SaPE2	PE3-150	pPC1000 edit site	pPC1000 edit site	+6 G to C	+50	GTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAGAGCCTGCACTGTAATGAAAATTACAGGTCTGATAGGCAGCCTGCACTGTAATGAAAATTACAG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGCT (SEQ ID NO: 5202)	10		GAGAACCGTC CTAACAAAT CTGC (SEQ ID NO: 5186)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG AACTTGTTGGCGAGA (SEQ ID NO: 5183)

46D	HeLa CRISPRi cells with integrated pPC1000 insert	Chen2021_0049	SaPE 2–P2 A–BS D	PE2	pPC1000 edit site	pPC1000 0 edit site	+6 G to C		TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG GAATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGCT (SEQ ID NO: 5202)	0	0	0	0	0
46D	HeLa CRISPRi cells with integrated pPC1000 insert	Chen2021_0050	SaPE 2–P2 A–BS D	PE2, +dasatinib	pPC1000 edit site	pPC1000 0 edit site	+6 G to C		TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGCT (SEQ ID NO: 5202)	0	0	0	0	0
46D	HeLa CRISPRi cells with integrated pPC1000 insert	Chen2021_0051	SaPE 2–P2 A–BS D	PE3+50	pPC1000 edit site	pPC1000 0 edit site	+6 G to C	+50	TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGCT (SEQ ID NO: 5202)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTCGTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46D	HeLa CRISPRi cells with integrated pPC1000 insert	Chen2021_0052	SaPE 2–P2 A–BS D	PE3+50, +dasatinib	pPC1000 edit site	pPC1000 0 edit site	+6 G to C	+50	TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGCT (SEQ ID NO: 5202)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTCGTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46D	HeLa CRISPRi cells with integrated pPC1000 insert	Chen2021_0053	SaPE 2–P2 A–BS D	PE3–50	pPC1000 edit site	pPC1000 0 edit site	+6 G to C	-50	TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGCT (SEQ ID NO: 5202)	0	0	GAGAACGTC CTAACAAAT CTGC (SEQ ID NO: 5188)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTCGTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
46D	HeLa CRISPRi cells with integrated pPC1000 insert	Chen2021_0054	SaPE 2–P2 A–BS D	PE3–50, +dasatinib	pPC1000 edit site	pPC1000 0 edit site	+6 G to C	-50	TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT CT (SEQ ID NO: 5201)	AGAATTCACCGCACGAGTG GCAGGCT (SEQ ID NO: 5202)	0	0	GAGAACGTC CTAACAAAT CTGC (SEQ ID NO: 5188)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTCGTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
54B	HeLa CRISPRi cells with integrated pPC1000 insert	Chen2021_0055	SaPE 2–P2 A–BS D	PE2, screen, pegRNA	pPC1000 edit site	pPC1000 0 edit site	+6 G to C		TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTCACCGCACGAGT GCAGGCT (SEQ ID NO: 5186)	0	0	0	0	0
54B	HeLa CRISPRi cells with integrated pPC1000 insert, MSH2	Chen2021_0056	SaPE 2–P2 A–BS D	PE2, pegRNA with decoded scaffold	pPC1000 edit site	pPC1000 0 edit site	+6 G to C		TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGGCATCTTTGATGTCGAGA (SEQ ID NO: 5228)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTCACCGCACGAGT GCAGGCT (SEQ ID NO: 5186)	0	0	0	0	0
54B	HeLa CRISPRi cells with integrated pPC1000 insert	Chen2021_0057	SaPE 2–P2 A–BS D	PE2, pegRNA with decoded scaffold	pPC1000 edit site	pPC1000 0 edit site	+6 G to C		TGTCTGATAGGC AGCCTGCACG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGGCATCTTTGATGTCGAGA (SEQ ID NO: 5228)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTCACCGCACGAGT GCAGGCT (SEQ ID NO: 5186)	0	0	0	0	0

54B	HeLa CR(SP)] cells with integrated pPC1000 (MS42	Chen2021 0059	SaPE 2–P2 A–BS D	PE2, pegRNA with nicking scaffold	pPC1000 edit site	pPC100 0 edit site	+6 G to C	-50	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TCGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTGACGCACCAGT GCAGGCT (SEQ ID NO: 5186)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
54B	HeLa CR(SP)] cells with integrated pPC1000 (nch	Chen2021 0059	SaPE 2–P2 A–BS D	PE3+50, screen pegRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to C	-50	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TCGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTGACGCACCAGT GCAGGCT (SEQ ID NO: 5186)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
54B	HeLa CR(SP)] cells with integrated pPC1000 (MS42	Chen2021 0060	SaPE 2–P2 A–BS D	PE3+50, screen pegRNA with nicking scaffold	pPC1000 edit site	pPC100 0 edit site	+6 G to C	-50	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5229)	AGCCTGCAC (SEQ ID NO: 5184)	TCGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTGACGCACCAGT GCAGGCT (SEQ ID NO: 5186)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
54B	HeLa CR(SP)] cells with integrated pPC1000 (MS42	Chen2021 0061	SaPE 2–P2 A–BS D	PE3+50, pegRNA with nicking scaffold	pPC1000 edit site	pPC100 0 edit site	+6 G to C	-50	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5229)	AGCCTGCAC (SEQ ID NO: 5184)	TCGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTGACGCACCAGT GCAGGCT (SEQ ID NO: 5186)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
54B	HeLa CR(SP)] cells with integrated pPC1000 (MS42	Chen2021 0062	SaPE 2–P2 A–BS D	PE3+50, pegRNA with nicking scaffold	pPC1000 edit site	pPC100 0 edit site	+6 G to C	-50	GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGGGATCTTCGATGTCGAGA (SEQ ID NO: 5229)	AGCCTGCAC (SEQ ID NO: 5184)	TCGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTGACGCACCAGT GCAGGCT (SEQ ID NO: 5186)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	0	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5183)
42C	HEK293 T	Chen2021 _0063	PE2	NT siRNA	HEK3	pPC1000 edit site	HEK3	+5 G to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
42C	HEK293 T	Chen2021 _0064	PE2	NSH2 siRNA	HEK3	pPC1000 edit site	HEK3	+5 G to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
42C	HEK293 T	Chen2021 _0065	PE2	NSH6 siRNA	HEK3	pPC1000 edit site	HEK3	+5 G to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
42C	HEK293 T	Chen2021 _0066	PE2	NL-H1 siRNA	HEK3	pPC1000 edit site	HEK3	+5 G to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
42C	HEK293 T	Chen2021 _0067	PE2	PMS2 siRNA	HEK3	pPC1000 edit site	HEK3	+5 G to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
42C	HEK293 T	Chen2021 _0068	PE2	NT siRNA	HEK3	pPC1000 edit site	HEK3	+5 G to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA (SEQ ID NO: 5234)	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
42C	HEK293 T	Chen2021 _0069	PE2	NSH2 siRNA	HEK3	pPC1000 edit site	HEK3	+5 G to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA (SEQ ID NO: 5234)	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

42C	HEK293T	Chen2021_0070	PE2	MSH6 siRNA	HEK3	+5 G to T	+90	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCA TC (SEQ ID NO: 5234)	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
42C	HEK293T	Chen2021_0071	PE2	MLH1 siRNA	HEK3	+5 G to T	+90	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCA TC (SEQ ID NO: 5234)	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
42C	HEK293T	Chen2021_0072	PE2	PMS2 siRNA	HEK3	+5 G to T	+90	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCA TC (SEQ ID NO: 5234)	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
42C	HEK293T	Chen2021_0073	PE2	NT siRNA	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	0	0
42C	HEK293T	Chen2021_0074	PE2	MSH2 siRNA	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	0	0
42C	HEK293T	Chen2021_0075	PE2	MSH6 siRNA	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	0	0
42C	HEK293T	Chen2021_0076	PE2	MLH1 siRNA	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	0	0
42C	HEK293T	Chen2021_0077	PE2	PMS2 siRNA	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	0	0
42C	HEK293T	Chen2021_0078	PE2	NT siRNA	EMX1	+5 G to T	-57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCGCAATC (SEQ ID NO: 5239)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	GCCGTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
42C	HEK293T	Chen2021_0079	PE2	MSH2 siRNA	EMX1	+5 G to T	-57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	GCCGTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
42C	HEK293T	Chen2021_0080	PE2	MSH6 siRNA	EMX1	+5 G to T	-57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	GCCGTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
42C	HEK293T	Chen2021_0081	PE2	MLH1 siRNA	EMX1	+5 G to T	-57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	GCCGTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
42C	HEK293T	Chen2021_0082	PE2	PMS2 siRNA	EMX1	+5 G to T	-57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCGCAATC (SEQ ID NO: 5240)	GTTGATGGGAGCAACTTCT (SEQ ID NO: 5240)	0	0	GCCGTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
42C	HEK293T	Chen2021_0083	PE2	NT siRNA	RUNX1	+3 A to C		GCATTTTCAGG AGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	ATTTTCAGGAGGA (SEQ ID NO: 5244)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCT CCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
42C	HEK293T	Chen2021_0084	PE2	NSH2 siRNA	RUNX1	+3 A to C		GCATTTTCAGG AGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	ATTTTCAGGAGGA AG (SEQ ID NO: 5244)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCT CCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0

42C	HEK293T	Chen2021_0035	PE2	MSH6 siRNA	RUNX1	+3 A to C		GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGCTGAAGCCAGGGCT (TGTCTGCTGAAAAT (SEQ ID NO: 5245)	19	0	0	0	0					
42C	HEK293T	Chen2021_0095	PE2	MLH1 siRNA	RUNX1	+3 A to C		GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGCTGAAGCCAGGGCT (TGTCTGCTGAAAAT (SEQ ID NO: 5245)	19	0	0	0	0					
42C	HEK293T	Chen2021_0097	PE2	FMS2 siRNA	RUNX1	+3 A to C		GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGCTGAAGCCAGGGCT (TGTCTGCTGAAAAT (SEQ ID NO: 5245)	19	0	0	0	0					
42C	HEK293T	Chen2021_0089	PE2	NT siRNA	RUNX1	+3 A to C	+38	GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGCTGAAGCCAGGGCT (TGTCTGCTGAAAAT (SEQ ID NO: 5245)	19	0	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)			
42C	HEK293T	Chen2021_0090	PE2	MSH2 siRNA	RUNX1	+3 A to C	+38	GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGCTGAAGCCAGGGCT (TGTCTGCTGAAAAT (SEQ ID NO: 5245)	19	0	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)			
42C	HEK293T	Chen2021_0091	PE2	MLH1 siRNA	RUNX1	+3 A to C	+38	GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGCTGAAGCCAGGGCT (TGTCTGCTGAAAAT (SEQ ID NO: 5245)	19	0	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)			
42C	HEK293T	Chen2021_0092	PE2	FMS2 siRNA	RUNX1	+3 A to C	+38	GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGCTGAAGCCAGGGCT (TGTCTGCTGAAAAT (SEQ ID NO: 5245)	19	0	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)			
42D	HAP1 wild-type	Chen2021_0093	PE2-P2A-BSD	HEK3	RUNX1	+5 G to T		GGGCCAGACTGAGCAGCGTGA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	CAGACTGAGCAGCAGG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCACT (GTCTG (SEQ ID NO: 5234)	19	0	0	0	0					
42D	HAP1 ΔMSH2	Chen2021_0094	PE2-P2A-BSD	HEK3	RUNX1	+5 G to T		GGGCCAGACTGAGCAGCGTGA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	CAGACTGAGCAGCAGG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCACT (GTCTG (SEQ ID NO: 5234)	19	0	0	0	0					
42D	HAP1 ΔMLH1	Chen2021_0095	PE2-P2A-BSD	HEK3	RUNX1	+5 G to T		GGGCCAGACTGAGCAGCGTGA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	CAGACTGAGCAGCAGG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCACT (GTCTG (SEQ ID NO: 5234)	19	0	0	0	0					
42D	HAP1 wild-type	Chen2021_0096	PE2-P2A-BSD	HEK3	RUNX1	+5 G to T	+90	GCGCCAGACTGAGCAGCGTGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5241)	CAGACTGAGCAGCAGG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCACT (GTCTG (SEQ ID NO: 5234)	19	0	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)			
42D	HAP1 ΔMSH2	Chen2021_0097	PE2-P2A-BSD	HEK3	RUNX1	+5 G to T	+90	GGGCCAGACTGAGCAGCGTGA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	CAGACTGAGCAGCAGG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCACT (GTCTG (SEQ ID NO: 5234)	19	0	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)			
42D	HAP1 ΔMLH1	Chen2021_0098	PE2-P2A-BSD	HEK3	RUNX1	+5 G to T	+90	GGGCCAGACTGAGCAGCGTGA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5239)	CAGACTGAGCAGCAGG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCACT (GTCTG (SEQ ID NO: 5234)	19	0	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)			
42D	HAP1 wild-type	Chen2021_0099	PE2-P2A-BSD	EMX1	EMX1	+5 G to T		GAGTGCGAGAGAGGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGAGAGAAGGA (SEQ ID NO: 5238)	GAAGTGCTCCCATC A (SEQ ID NO: 5239)	GTGATGGGAACACTTCT (TCTTC TGCTCCGGA (SEQ ID NO: 5240)	19	0	0	0	0					

42D	HAP1 ΔMSH2	Chen2021_0100	PE2-P2A-BSD	PE2	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0						
42D	HAP1 ΔMLH1	Chen2021_0101	PE2-P2A-BSD	PE2	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0						
42D	HAP1 wild-type	Chen2021_0102	PE2-P2A-BSD	PE3	EMX1	+5 G to T	57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)				
42D	HAP1 ΔMSH2	Chen2021_0103	PE2-P2A-BSD	PE3	EMX1	+5 G to T	57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)				
42D	HAP1 ΔMLH1	Chen2021_0104	PE2-P2A-BSD	PE3	EMX1	+5 G to T	57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)				
48D	HEK293 T	Chen2021_0105	PE2	NT siRNA, no pre-treatment	HEK3	+5 G to T		GCGCCAGAACTG AGCACGGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACG TGAATTGCAGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0						
48D	HEK293 T	Chen2021_0106	PE2	MLH1 siRNA, no pre-treatment	HEK3	+5 G to T		GCGCCAGAACTG AGCACGGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACG TGAATTGCAGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0						
48D	HEK293 T	Chen2021_0107	PE2	NT siRNA no pre-treatment	HEK3	+5 G to T	+90	GCGCCAGAACTG AGCACGGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACG TGAATTGCAGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	GTCAACGAG TATCCCGGT GC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)				
48D	HEK293 T	Chen2021_0108	PE2	MLH1 siRNA, no pre-treatment	HEK3	+5 G to T	+90	GCGCCAGAACTG AGCACGGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACG TGAATTGCAGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	GTCAACGAG TATCCCGGT GC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)				
48D	HEK293 T	Chen2021_0109	PE2	NT siRNA, no pre-treatment	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0						
48D	HEK293 T	Chen2021_0110	PE2	MLH1 siRNA, no pre-treatment	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0						
48D	HEK293 T	Chen2021_0111	PE2	NT siRNA no pre-treatment	EMX1	+5 G to T	57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)				
48D	HEK293 T	Chen2021_0112	PE2	MLH1 siRNA, no pre-treatment	EMX1	+5 G to T	57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA A (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0						
48D	HEK293 T	Chen2021_0113	PE2	NT siRNA no pre-treatment	RUNX1	+5 G to T		GCATTTTCAGGG AGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	AATTTTCAGGAGGA A (SEQ ID NO: 5244)	CGGCTGGCTTCAGAC TGCTGCTGAAAAT (SEQ ID NO: 5245)	TCGTCTGAAGCCAGGGACT TCGTCGCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0						
48D	HEK293 T	Chen2021_0114	PE2	MLH1 siRNA no pre-treatment	RUNX1	+5 G to T		GCATTTTCAGGG AGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	AATTTTCAGGAGGA A (SEQ ID NO: 5244)	CGGCTGGCTTCAGAC TGCTGCTGAAAAT (SEQ ID NO: 5245)	TCGTCTGAAGCCAGGGACT TCGTCGCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0						

48D	HEK293T	Chen2021_0115	PE2	NT siRNA, no pre-treatment	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
48D	HEK293T	Chen2021_0116	PE2	MLH1 siRNA, no pre-treatment	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43A,43C	HEK293T	Chen2021_0117	PE2	Control	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A	HEK293T	Chen2021_0118	PE2	MSH2	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A	HEK293T	Chen2021_0119	PE2	MSH2 K675R	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A	HEK293T	Chen2021_0120	PE2	MSH6	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A	HEK293T	Chen2021_0121	PE2	MSH6 K1140R	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A	HEK293T	Chen2021_0122	PE2	PMS2	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A	HEK293T	Chen2021_0123	PE2	PMS2 E41A	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A	HEK293T	Chen2021_0124	PE2	PMS2 E705K	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A,43C	HEK293T	Chen2021_0125	PE2	MLH1	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A,43C	HEK293T	Chen2021_0126	PE2	MLH1 E34A	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A,43C	HEK293T	Chen2021_0127	PE2	MLH1 Δ756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43C	HEK293T	Chen2021_0128	PE2	MLH1 Δ754-756 (MLH1dn)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43C	HEK293T	Chen2021_0129	PE2	MLH1 E34A Δ754-756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43C	HEK293T	Chen2021_0130	PE2	MLH1 1-335	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0

43C	HEK293T	Chen2021_0131	PE2	MLH1 1-335 E34A	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43C	HEK293T	Chen2021_0132	PE2	MLH1 1-335-bpSV40NLS GS codon (MLH1NTD-NLS)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43C	HEK293T	Chen2021_0133	PE2	MLH1 501-756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43C	HEK293T	Chen2021_0134	PE2	MLH1 Δ2-500 Δ754-756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43C	HEK293T	Chen2021_0135	PE2	MLH1 Δ2-460 Δ754-756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43C	HEK293T	Chen2021_0136	PE2	bpSV40NLS–MLH1 Δ2-500 Δ754-756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43A, 43C	HEK293T	Chen2021_0137	PE2	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A	HEK293T	Chen2021_0138	PE2	MSH2	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A	HEK293T	Chen2021_0139	PE2	MSH2 K675R	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A	HEK293T	Chen2021_0140	PE2	MSH6	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A	HEK293T	Chen2021_0141	PE2	MSH6 K1140R	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A	HEK293T	Chen2021_0142	PE2	PMS2	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A	HEK293T	Chen2021_0143	PE2	PMS2 E41A	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A	HEK293T	Chen2021_0144	PE2	PMS2 E705K	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A, 43C	HEK293T	Chen2021_0145	PE2	MLH1	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A, 43C	HEK293T	Chen2021_0146	PE2	MLH1 E34A	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0

43A,43C	HEK293T	Chen2021_0147	PE2	MLH1 Δ756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0148	PE2	MLH1 Δ754-756 (MLH1dn)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0149	PE2	MLH1 E34A Δ754-756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0150	PE2	MLH1 1-335	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0151	PE2	MLH1 1-335 E34A	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0152	PE2	MLH1 1-335–bpSV40NLS GS codon (MLH1NTD–NLS)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0153	PE2	MLH1 501-756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0154	PE2	MLH1 Δ2-500 Δ754-756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0155	PE2	MLH1 Δ2-460 Δ754-756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43C	HEK293T	Chen2021_0156	PE2	bpSV40NLS–MLH1 Δ2-500 Δ754-756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43A,43C	HEK293T	Chen2021_0157	PE2	Control	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A	HEK293T	Chen2021_0158	PE2	MSH2	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A	HEK293T	Chen2021_0159	PE2	MSH2 K675R	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A	HEK293T	Chen2021_0160	PE2	MSH6	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A	HEK293T	Chen2021_0161	PE2	MSH6 K1140R	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A	HEK293T	Chen2021_0162	PE2	PMS2	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0

43A	HEK293T	Chen2021_0163	PE2	PMS2 E41A	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A	HEK293T	Chen2021_0164	PE2	PMS2 E705K	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A, 43C	HEK293T	Chen2021_0165	PE2	MLH1	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A, 43C	HEK293T	Chen2021_0166	PE2	MLH1 E34A	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43A, 43C	HEK293T	Chen2021_0167	PE2	MLH1 Δ756	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0168	PE2	MLH1 Δ754-756 (MLH1dn)	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0169	PE2	MLH1 E34A Δ754-756	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0170	PE2	MLH1 1-335	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0171	PE2	MLH1 1-335 E34A	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0172	PE2	MLH1 1-335-bpSV40NLS GS codon (MLH1NTD-NLS)	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0173	PE2	MLH1 501-756	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0174	PE2	MLH1 Δ2-500 Δ754-756	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0175	PE2	MLH1 Δ2-460 Δ754-756	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43C	HEK293T	Chen2021_0176	PE2	bpSV40NLS–MLH1 Δ2-500 Δ754-756	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43D	HEK293T	Chen2021_0177	PE2	Control	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43D	HEK293T	Chen2021_0178	PE2	MLH1 Δ754-756 (MLH1dn)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0

43D	HEK293T	Chen2021_0179	PE2	MLH1 1-335-bpSV40NLS GS codon (MLH1NTD-NLS)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43D	HEK293T	Chen2021_0180	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
43D	HEK293T	Chen2021_0181	PE2	Control	HEK3	HEK3	+5 G to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0182	PE2	MLH1 Δ754-756 (MLH1dn)	HEK3	HEK3	+5 G to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0183	PE2	MLH1 1-335-bpSV40NLS GS codon (MLH1NTD-NLS)	HEK3	HEK3	+5 G to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0184	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	HEK3	HEK3	+5 G to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0185	PE2	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43D	HEK293T	Chen2021_0186	PE2	MLH1 Δ754-756 (MLH1dn)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43D	HEK293T	Chen2021_0187	PE2	MLH1 1-335-bpSV40NLS GS codon (MLH1NTD-NLS)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43D	HEK293T	Chen2021_0188	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
43D	HEK293T	Chen2021_0189	PE2	Control	EMX1	EMX1	+5 G to T	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0190	PE2	MLH1 Δ754-756 (MLH1dn)	EMX1	EMX1	+5 G to T	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0191	PE2	MLH1 1-335-bpSV40NLS GS codon (MLH1NTD-NLS)	EMX1	EMX1	+5 G to T	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

43D	HEK293T	Chen2021_0192	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	EMX1	EMX1	+5 G to T	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0193	PE2	Control	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43D	HEK293T	Chen2021_0194	PE2	MLH1 Δ754-756 (MLH1dn)	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43D	HEK293T	Chen2021_0195	PE2	MLH1 1-335–bpSV40NLS GS codon (MLH1NTD–NLS)	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43D	HEK293T	Chen2021_0196	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
43D	HEK293T	Chen2021_0197	PE2	Control	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0198	PE2	MLH1 Δ754-756 (MLH1dn)	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0199	PE2	MLH1 1-335–bpSV40NLS GS codon (MLH1NTD–NLS)	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43D	HEK293T	Chen2021_0200	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0201	PE2	PE2	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
43F	HEK293T	Chen2021_0202	PE2	PE2	EMX1	EMX1	+5 G to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	0	0
43F	HEK293T	Chen2021_0203	PE2	PE2	RNF2	RNF2	+1 C to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TTGAGGTGTTCGTT (SEQ ID NO: 5253)	AACGAACACCTCAAGTAATGACTAAGATG (SEQ ID NO: 5254)	0	0	0	0
43F	HEK293T	Chen2021_0204	PE2	PE2	FANCF	FANCF	+4 T to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCAGGATCGCTTTTCC (SEQ ID NO: 5257)	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT (SEQ ID NO: 5258)	0	0	0	0
43F	HEK293T	Chen2021_0205	PE2	PE2	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
43F	HEK293T	Chen2021_0206	PE2	PE2	DNMT1	DNMT1	+8 T to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGCGACGGGAGGGC (SEQ ID NO: 5263)	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5264)	0	0	0	0

43F	HEK293T	Chen2021_0207	PE2	PE2	VEGFA	VEGFA	+3 A to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGTGGGCTCCAGATGGCACATT (SEQ ID NO: 5267)	AATGTGCCATCTGGAGCCCACATCTGGCCTGCAGA (SEQ ID NO: 5268)	0	0	0	0
43F	HEK293T	Chen2021_0208	PE2	PE4	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
43F	HEK293T	Chen2021_0209	PE2	PE4	EMX1	EMX1	+5 G to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	0	0
43F	HEK293T	Chen2021_0210	PE2	PE4	RNF2	RNF2	+1 C to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TTGAGGTGTTCGTT (SEQ ID NO: 5253)	AACGAACACCTCAAGTAATGACTAAGATG (SEQ ID NO: 5254)	0	0	0	0
43F	HEK293T	Chen2021_0211	PE2	PE4	FANCF	FANCF	+4 T to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCAGGATCGCTTTTCC (SEQ ID NO: 5257)	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT (SEQ ID NO: 5258)	0	0	0	0
43F	HEK293T	Chen2021_0212	PE2	PE4	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
43F	HEK293T	Chen2021_0213	PE2	PE4	DNMT1	DNMT1	+8 T to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGCGACGGGAGGGC (SEQ ID NO: 5263)	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5264)	0	0	0	0
43F	HEK293T	Chen2021_0214	PE2	PE4	VEGFA	VEGFA	+3 A to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGTGGGCTCCAGATGGCACATT (SEQ ID NO: 5267)	AATGTGCCATCTGGAGCCCACATCTGGCCTGCAGA (SEQ ID NO: 5268)	0	0	0	0
43F	HEK293T	Chen2021_0215	PE2	PE3	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0216	PE2	PE3	EMX1	EMX1	+5 G to A	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0217	PE2	PE3	RNF2	RNF2	+1 C to T	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TTGAGGTGTTCGTT (SEQ ID NO: 5253)	AACGAACACCTCAAGTAATGACTAAGATG (SEQ ID NO: 5254)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0218	PE2	PE3	FANCF	FANCF	+4 T to A	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCAGGATCGCTTTTCC (SEQ ID NO: 5257)	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT (SEQ ID NO: 5258)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0219	PE2	PE3	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0220	PE2	PE3	DNMT1	DNMT1	+8 T to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGCGACGGGAGGGC (SEQ ID NO: 5263)	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5264)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0221	PE2	PE3	VEGFA	VEGFA	+3 A to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGTGGGCTCCAGATGGCACATT (SEQ ID NO: 5267)	AATGTGCCATCTGGAGCCCACATCTGGCCTGCAGA (SEQ ID NO: 5268)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

43F	HEK293T	Chen2021_0222	PE2	PE5	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0223	PE2	PE5	EMX1	EMX1	+5 G to A	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0224	PE2	PE5	RNF2	RNF2	+1 C to T	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TTGAGGTGTTCGTT (SEQ ID NO: 5253)	AACGAACACCTCAAGTAATGACTAAGATG (SEQ ID NO: 5254)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0225	PE2	PE5	FANCF	FANCF	+4 T to A	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCAGGATCGCTTTTCC (SEQ ID NO: 5257)	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT (SEQ ID NO: 5258)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0226	PE2	PE5	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0227	PE2	PE5	DNMT1	DNMT1	+8 T to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGCGACGGGAGGGC (SEQ ID NO: 5263)	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5264)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
43F	HEK293T	Chen2021_0228	PE2	PE5	VEGFA	VEGFA	+3 A to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGTGGGCTCCAGATGGCACATT (SEQ ID NO: 5267)	AATGTGCCATCTGGAGCCCACATCTGGCCTGCAGA (SEQ ID NO: 5268)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49A	HEK293T	Chen2021_0229	PE2	Control	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0230	PE2	MLH1 Δ754-756 (MLH1dn)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0231	PE2	PMS2 Δ2-607 E705K	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0232	PE2	PMS2 Δ2-635 E705K	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0233	PE2	NLS-PMS2 Δ1-635 E705K	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0234	PE2	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0235	PE2	MLH1 Δ754-756 (MLH1dn)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0236	PE2	PMS2 Δ2-607 E705K	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0

49A	HEK293T	Chen2021_0237	PE2	PMS2 Δ2-636 E705K	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCGGAGCGAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGACGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0238	PE2	NLS-PMS2-ΔN-636-E705K	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCGAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGACGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0239	PE2	Control	RUNX1	+3 A to C		GCATTTTCAGGGAAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	ATTTTCAGGAGGGA (SEQ ID NO: 5244)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0240	PE2	ML-H1 Δ754-756 (ML-H1dn)	RUNX1	+3 A to C		GCATTTTCAGGGAAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	ATTTTCAGGAGGGA (SEQ ID NO: 5244)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0241	PE2	PMS2 Δ2-607 E705K	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	ATTTTCAGGAGGGA (SEQ ID NO: 5244)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0242	PE2	PMS2 Δ2-636 E705K	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	ATTTTCAGGAGGGA (SEQ ID NO: 5244)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0243	PE2	NLS-PMS2-ΔN-636-E705K	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	ATTTTCAGGAGGGA (SEQ ID NO: 5244)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0244	PE2	Control	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0245	PE2	ML-H1 E34A	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0246	PE2	ML-H1 Δ756	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0247	PE2	ML-H1 Δ754-756 (ML-H1dn)	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0248	PE2	ML-H1 E34A Δ754-756	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0249	PE2	PMS2 1-134	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0250	PE2	PMS2 1-134 E41A	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0251	PE2-P2A-ML-H1 Δ754-756	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0252	ML-H1 Δ754-756--P2A-PE2	HEK3	+5 G to T		GCGCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACGTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAG TGTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0	0	0

49A	HEK293T	Chen2021_0253	ML-H1 Δ754-756–PE2	ML-H1 Δ754-756–PE2	HEK3	+5 G to T	GGCCCAGACTG (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGAGTGAGCACG (SEQ ID NO: 5232)	TGATTCGAGA (SEQ ID NO: 5233)	TCTTCGAATCAGGTGCTCAA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0254	PE2	Control	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0255	PE2	ML-H1 E34A	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0256	PE2	ML-H1 Δ756	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0257	PE2	ML-H1 Δ754-756 (ML-H1Δ9)	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0259	PE2	ML-H1 E34A Δ754-756	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0259	PE2	FMS2 -134	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0260	PE2	FMS2 -134 E41A	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0261	PE2–P2A–ML-H1 Δ754-756	PE2–P2A–ML-H1 Δ754-756	EMX1	+5 G to T	CAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0262	ML-H1 Δ754-756–P2A–PE2	ML-H1 Δ754-756–P2A–PE2	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0263	ML-H1 Δ754-756–PE2	ML-H1 Δ754-756–PE2	EMX1	+5 G to T	GAGTCCGAGCA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0264	PE2	Control	RUNX1	+3 A to C	GCATTTTCAGG (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGGCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0265	PE2	ML-H1 E34A	RUNX1	+3 A to C	GCATTTTCAGG (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGGCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0266	PE2	ML-H1 Δ756	RUNX1	+3 A to C	GCATTTTCAGG (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGGCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0267	PE2	ML-H1 Δ754-756 (ML-H1Δ9)	RUNX1	+3 A to C	GCATTTTCAGG (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGGCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	16	0	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0268	PE2	ML-H1 E34A Δ754-756	RUNX1	+3 A to C	GCATTTTCAGG (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA (SEQ ID NO: 5243)	CGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGGCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	16	0	0	0	0	0	0	0	0

49A	HEK293T	Chen2021_0269	PE2	PMS2 1-134	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)		0	0	0
49A	HEK293T	Chen2021_0270	PE2	PMS2 1-134 E41A	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)		0	0	0
49A	HEK293T	Chen2021_0271	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)		0	0	0
49A	HEK293T	Chen2021_0272	MLH1 Δ754-756–P2A–PE2	MLH1 Δ754-756–P2A–PE2	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)		0	0	0
49A	HEK293T	Chen2021_0273	MLH1 Δ754-756–PE2	MLH1 Δ754-756–PE2	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)		0	0	0
49A	HEK293T	Chen2021_0274	PE2	Control	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0275	PE2	MLH1 Δ754-756 (MLH1dn)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0276	PE2	MSH6 Δ2-361	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0277	PE2	MSH6 Δ2-361 K1140R	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0278	PE2	CDKN1A	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0279	PE2	MLH1 1-335	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0280	PE2	MLH1 1-226	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0281	PE2	MLH1 1-210	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0282	PE2	MLH1 1-335 E34A	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0283	PE2	MLH1 1-226 E34A	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0
49A	HEK293T	Chen2021_0284	PE2	MLH1 1-210 E34A	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)		0	0	0

49A	HEK293T	Chen2021_0285	PE2	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0286	PE2	MLH1 Δ754-756 (MLH1dn)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0287	PE2	MSH6 Δ2-361	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0288	PE2	MSH6 Δ2-361 K1140R	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0289	PE2	CDKN1A	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0290	PE2	MLH1 1-335	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0291	PE2	MLH1 1-226	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0292	PE2	MLH1 1-210	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0293	PE2	MLH1 1-335 E34A	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0294	PE2	MLH1 1-226 E34A	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0295	PE2	MLH1 1-210 E34A	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0296	PE2	Control	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0297	PE2	MLH1 Δ754-756 (MLH1dn)	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0298	PE2	MSH6 Δ2-361	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0299	PE2	MSH6 Δ2-361 K1140R	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0300	PE2	CDKN1A	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0301	PE2	MLH1 1-335	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0

49A	HEK293T	Chen2021_0302	PE2	MLH1 1-226	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0303	PE2	MLH1 1-210	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0304	PE2	MLH1 1-335 E34A	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0305	PE2	MLH1 1-226 E34A	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0306	PE2	MLH1 1-210 E34A	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49A	HEK293T	Chen2021_0307	PE2	Control	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0308	PE2	MLH1 Δ754-756 (MLH1dn)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0309	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0310	PE2–bpSV40NLS–MLH1 Δ754-756	PE2–bpSV40NLS–MLH1 Δ754-756	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0311	PE2–MLH1 Δ754-756–bpSV40NLS	PE2–MLH1 Δ754-756–bpSV40NLS	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0312	PE2	bpSV40NLS–MLH1 1-335	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0313	PE2	MLH1 1-335–bpSV40NLS	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0314	PE2	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0315	PE2	MLH1 Δ754-756 (MLH1dn)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0316	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0

49A	HEK293T	Chen2021_0317	PE2-bpSV40NLS-MLH1 Δ754-756	EMX1	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGACGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0318	PE2-MLH1 Δ754-756-bpSV40NLS	EMX1	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGACGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0319	PE2-bpSV40NLS-MLH1-335	EMX1	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGACGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0326	PE2-MLH1-335-bpSV40NLS	EMX1	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGAA (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGACGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0321	PE2 Control	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0322	PE2 MLH1 Δ754-756 (MLH1nh)	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0323	PE2 MLH1 Δ754-756 GS codon (MLH1nh)	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0324	PE2-P2A-MLH1 Δ754-756	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0325	PE2-P2A-MLH1 Δ754-756 GS codon	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0326	PE2 MLH1-335-bpSV40NLS	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0327	PE2 MLH1-335-bpSV40NLS GS codon	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0328	PE2 MLH1-400 GS codon	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0329	PE2 MLH1-460 GS codon	HEK3	+5 G to T	GGGCCGAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0330	PE2 MLH1-485 GS codon	HEK3	+5 G to T	CTCCCCAGACGTGCGTTTAT (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0

49A	HEK293T	Chen2021_0331	PE2	MLH1 1-501 GS codon	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0332	PE2	MLH1 1-400–bpSV40NLS GS codon	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0333	PE2	MLH1 1-460–bpSV40NLS GS codon	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0334	PE2	MLH1 1-485–bpSV40NLS GS codon	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0335	PE2	MLH1 1-501–bpSV40NLS GS codon	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49A	HEK293T	Chen2021_0336	PE2	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0337	PE2	MLH1 Δ754-756 (MLH1dn)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0338	PE2	MLH1 Δ754-756 GS codon (MLH1dn)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0339	PE2–P2A–MLH1 Δ754-756	PE2–P2A–MLH1 Δ754-756	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0340	PE2–P2A–MLH1 Δ754-756 (GS codon)	PE2–P2A–MLH1 Δ754-756 GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0341	PE2	MLH1 1-335–bpSV40NLS	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0342	PE2	MLH1 1-335–bpSV40NLS GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0343	PE2	MLH1 1-400 GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0344	PE2	MLH1 1-460 GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0345	PE2	MLH1 1-485 GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49A	HEK293T	Chen2021_0346	PE2	MLH1 1-501 GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0

49A	HEK293T	Chen2021_0347	PE2	ML-H1 1-400--bpSV 40NLS GS codon	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0348	PE2	ML-H1 1-460--bpSV 40NLS GS codon	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0349	PE2	ML-H1 1-495--bpSV 40NLS GS codon	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0350	PE2	ML-H1 1-501--bpSV 40NLS GS codon	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	0	0	0	0	0	0	0
49A	HEK293T	Chen2021_0351	PE2	Control	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0352	PE2	ML-H1 Δ754-756 (ML-H1Δ)	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0353	PE2	ML-H1 Δ754-756 GS codon (ML-H1Δ1)	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0354	PE2	ML-H1 1-305--bpSV 40NLS GS codon	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0355	PE2-P2A-ML-H1 Δ754-756 GS codon	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0356	PE2-P2A-ML-H1 1-335 GS codon	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0357	ML-H1 Δ754-756 GS codon-PE2	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0358	PE2-ML-H1 Δ754-756 GS codon	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0359	ML-H1 1-335 GS codon-PE2	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0360	PE2-ML-H1 1-335 GS codon	HEK3	+5 G to T		GTCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGT (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0	0	0	0

49A	HEK293T	Chen2021_0361	PE2	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0362	PE2	ML-H1 Δ754-756 (MLH1del)	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0363	PE2	ML-H1 Δ754-756 GS codon (MLH1del)	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0364	PE2	ML-H1 1-335-pgsV 40NLS GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0365	PE2-P2A-ML-H1 Δ754-756 GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0366	PE2-P2A-ML-H1 1-335 GS codon GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0367	ML-H1 Δ754-756 GS codon-PE2	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0368	PE2-ML-H1 Δ754-756 GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0369	ML-H1 1-335 GS codon-PE2	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0370	PE2-ML-H1 1-335 GS codon	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA A (SEQ ID NO: 5238)	GAAGTGGTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0
49A	HEK293T	Chen2021_0371	Control	RUNX1	RUNX1	+3 A to C		GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5243)	AATTTTCAGGAGGA A (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC AG (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCGCTGCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0
49A	HEK293T	Chen2021_0372	PE2	ML-H1 Δ754-756 (MLH1del)	RUNX1	RUNX1	+3 A to C		GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGA A (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC AG (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCGCTGCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0
49A	HEK293T	Chen2021_0373	PE2	ML-H1 Δ754-756 GS codon (MLH1del)	RUNX1	RUNX1	+3 A to C		GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGA A (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC AG (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCGCTGCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0
49A	HEK293T	Chen2021_0374	PE2	ML-H1 1-335-pgsV 40NLS GS codon	RUNX1	RUNX1	+3 A to C		GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGA A (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC AG (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCGCTGCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0

49A	HEK293T	Chen2021_0375	PE2	PE2–P2A–ML-H1 Δ754-756 GS codon	RUNX1	+3 A to C	GCATTTTCAGGAGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAGA (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT (SEQ ID NO: 5245)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0376	PE2–P2A–ML-H1 1-335 GS codon	RUNX1	+3 A to C	GCATTTTCAGGAGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAGA (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT (SEQ ID NO: 5245)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0377	PE2–P2A–ML-H1 1-335 GS codon–PE2	RUNX1	+3 A to C	GCATTTTCAGGAGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAGA (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT (SEQ ID NO: 5245)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0378	ML-H1 Δ754-756 GS codon	PE2–ML-H1 Δ754-756 GS codon–PE2	RUNX1	+3 A to C	GCATTTTCAGGAGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAGA (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT (SEQ ID NO: 5245)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0379	ML-H1 1-335 GS Δ754-756 GS codon	PE2–ML-H1 1-335 GS codon–PE2	RUNX1	+3 A to C	GCATTTTCAGGAGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAGA (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT (SEQ ID NO: 5245)	0	0	0	0	0	0
49A	HEK293T	Chen2021_0380	ML-H1 1-335 GS codon–PE2	PE2–ML-H1 1-335 GS codon	RUNX1	+3 A to C	GCATTTTCAGGAGGAAGGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAGA (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT (SEQ ID NO: 5245)	0	0	0	0	0	0
49B	HEK293T	Chen2021_0381	PE2 (200 ng)	ML-H1sn (100 ng)	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGACACG (SEQ ID NO: 5232)	TGATTGCAGGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCACA (SEQ ID NO: 5234)	0	0	0	0	0	0
49B	HEK293T	Chen2021_0382	PE2 (200 ng)	ML-H1sn (50 ng)	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGACACG (SEQ ID NO: 5232)	TGATTGCAGGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCACA (SEQ ID NO: 5234)	0	0	0	0	0	0
49B	HEK293T	Chen2021_0383	PE2 (200 ng)	ML-H1sn (25 ng)	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGACACG (SEQ ID NO: 5232)	TGATTGCAGGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCACA (SEQ ID NO: 5234)	0	0	0	0	0	0
49B	HEK293T	Chen2021_0384	PE2 (200 ng)	ML-H1sn (12.5 ng)	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGACACG (SEQ ID NO: 5232)	TGATTGCAGGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCACA (SEQ ID NO: 5234)	0	0	0	0	0	0
49B	HEK293T	Chen2021_0385	PE2 (200 ng)	ML-H1sn (6.25 ng)	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGACACG (SEQ ID NO: 5232)	TGATTGCAGGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCACA (SEQ ID NO: 5234)	0	0	0	0	0	0
49B	HEK293T	Chen2021_0386	PE2 (200 ng)	ML-H1sn (3.125 ng)	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGACACG (SEQ ID NO: 5232)	TGATTGCAGGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCACA (SEQ ID NO: 5234)	0	0	0	0	0	0
49B	HEK293T	Chen2021_0387	PE2 (200 ng)	ML-H1sn (1.5625 ng)	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGACACG (SEQ ID NO: 5232)	TGATTGCAGGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCACA (SEQ ID NO: 5234)	0	0	0	0	0	0
49B	HEK293T	Chen2021_0388	PE2 (200 ng)	ML-H1sn (0.78125 ng)	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGACACG (SEQ ID NO: 5232)	TGATTGCAGGA (SEQ ID NO: 5233)	TCTGCAATCAGGTGCTCACA (SEQ ID NO: 5234)	0	0	0	0	0	0

49B	HEK293T	Chen2021_0389	PE2 (200 ng)	MLH1dn (0.390625 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0390	PE2 (200 ng)	MLH1dn (0 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0391	PE2 (200 ng)	MLH1dn (100 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0392	PE2 (200 ng)	MLH1dn (50 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0393	PE2 (200 ng)	MLH1dn (25 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0394	PE2 (200 ng)	MLH1dn (12.5 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0395	PE2 (200 ng)	MLH1dn (6.25 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0396	PE2 (200 ng)	MLH1dn (3.125 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0397	PE2 (200 ng)	MLH1dn (1.5625 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0398	PE2 (200 ng)	MLH1dn (0.78125 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0399	PE2 (200 ng)	MLH1dn (0.390625 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0400	PE2 (200 ng)	MLH1dn (0 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0401	PE2 (200 ng)	RFP (100 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0402	PE2 (100 ng)	RFP (50 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0403	PE2 (50 ng)	RFP (25 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0404	PE2 (25 ng)	RFP (12.5 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0405	PE2 (12.5 ng)	RFP (6.25 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0

49B	HEK293T	Chen2021_0406	PE2 (6.25 ng)	RFP (3.125 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0407	PE2 (3.125 ng)	RFP (1.5625 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0408	PE2 (1.5625 ng)	RFP (0.78125 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0409	PE2 (0.78125 ng)	RFP (0.390625 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0410	PE2 (0 ng)	RFP (0 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0411	PE2 (200 ng)	MLH1dn (100 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0412	PE2 (100 ng)	MLH1dn (50 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0413	PE2 (50 ng)	MLH1dn (25 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0414	PE2 (25 ng)	MLH1dn (12.5 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0415	PE2 (12.5 ng)	MLH1dn (6.25 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0416	PE2 (6.25 ng)	MLH1dn (3.125 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0417	PE2 (3.125 ng)	MLH1dn (1.5625 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0418	PE2 (1.5625 ng)	MLH1dn (0.78125 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0419	PE2 (0.78125 ng)	MLH1dn (0.390625 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0420	PE2 (0 ng)	MLH1dn (0 ng)	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
49B	HEK293T	Chen2021_0421	PE2 (200 ng)	RFP (100 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0422	PE2 (100 ng)	RFP (50 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0

49B	HEK293T	Chen2021_0423	PE2 (50 ng)	RFP (25 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0424	PE2 (25 ng)	RFP (12.5 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0425	PE2 (12.5 ng)	RFP (6.25 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0426	PE2 (6.25 ng)	RFP (3.125 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0427	PE2 (3.125 ng)	RFP (1.5625 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0428	PE2 (1.5625 ng)	RFP (0.78125 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0429	PE2 (0.78125 ng)	RFP (0.390625 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0430	PE2 (0 ng)	RFP (0 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0431	PE2 (200 ng)	MLH1dn (100 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0432	PE2 (100 ng)	MLH1dn (50 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0433	PE2 (50 ng)	MLH1dn (25 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0434	PE2 (25 ng)	MLH1dn (12.5 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0435	PE2 (12.5 ng)	MLH1dn (6.25 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0436	PE2 (6.25 ng)	MLH1dn (3.125 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0437	PE2 (3.125 ng)	MLH1dn (1.5625 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0438	PE2 (1.5625 ng)	MLH1dn (0.78125 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
49B	HEK293T	Chen2021_0439	PE2 (0.78125 ng)	MLH1dn (0.390625 ng)	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0

49B	HEK293T	Chen2021_0440	PE2 (0 ng)	ML-H1(in g)	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCGGAGCAGAAGA AGAGCTGAGCACG (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49C	HCT116	Chen2021_0441	PE2-P2A-BSD	Control	HEK3	HEK3	+5 G to T		GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTCGTCA (SEQ ID NO: 5234)	0	0	0	0	0	0
49C	HCT116	Chen2021_0442	PE2-P2A-BSD	ML-H1(in	HEK3	HEK3	+5 G to T		GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTCGTCA (SEQ ID NO: 5234)	0	0	0	0	0	0
49C	HCT116	Chen2021_0443	PE2-P2A-BSD	Control	HEK3	HEK3	+5 G to T	+90	GGGCCCAGACTG AGCACGTGA (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTCGTCA (SEQ ID NO: 5234)	0	0	0	GTCAACCAAG TATCCCGGT GC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
49C	HCT116	Chen2021_0444	PE2-P2A-BSD	ML-H1(in	HEK3	HEK3	+5 G to T	+90	GGGCCCAGACTG AGCACGTGA (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTCGCAGA (SEQ ID NO: 5233)	TCGTGCAATCACGTCGTCA (SEQ ID NO: 5234)	0	0	0	GTCAACCAAG TATCCCGGT GC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
49C	HCT116	Chen2021_0445	PE2-P2A-BSD	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA AGAGCTGAGCACG (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49C	HCT116	Chen2021_0446	PE2-P2A-BSD	ML-H1(in	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA AGAGCTGAGCACG (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49C	HCT116	Chen2021_0447	PE2-P2A-BSD	Control	EMX1	EMX1	+5 G to T	57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA AGAGCTGAGCACG (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	GCCGTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	0	0	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
49C	HCT116	Chen2021_0448	PE2-P2A-BSD	ML-H1(in	EMX1	EMX1	+5 G to T	57	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA AGAGCTGAGCACG (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	GCCGTTTGT ACTTTGTCC TC (SEQ ID NO: 5241)	0	0	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
49D	HEK293T	Chen2021_0449	PE2	Control	HEK3	HEK3	+5 G to T		CGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTCGTCA (SEQ ID NO: 5234)	0	0	0	0	0	0
49D	HEK293T	Chen2021_0450	PE2	mML-H1 Δ754-756 (human ML-H1(in	HEK3	HEK3	+5 G to T		CGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	GAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTCGTCA (SEQ ID NO: 5234)	0	0	0	0	0	0
49D	HEK293T	Chen2021_0451	PE2	mML-H1 757 GS codon (mouse	HEK3	HEK3	+5 G to T		CGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTCGTCA (SEQ ID NO: 5234)	0	0	0	0	0	0
49D	HEK293T	Chen2021_0452	PE2	Control	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA AGAGCTGAGCACG (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49D	HEK293T	Chen2021_0453	PE2	mML-H1 Δ754-756 (human ML-H1(in	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA AGAGCTGAGCACG (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49D	HEK293T	Chen2021_0454	PE2	mML-H1 757 GS codon (mouse	EMX1	EMX1	+5 G to T		GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA AGAGCTGAGCACG (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT TCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0	0	0
49D	HEK293T	Chen2021_0455	PE2	Control	RUNX1	RUNX1	+3 A to C		GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AAGAG (SEQ ID NO: 5243)	CGCTGGCTTCAGAAC A (SEQ ID NO: 5244)	TGTCTGCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0	0	0

49D	HEK293T	Chen2021_0456	PE2	hMLH1 Δ754-756 (human MLH1dn)	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49D	HEK293T	Chen2021_0457	PE2	mMLH1 1-757 GS codon (mouse	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
49E	N2A	Chen2021_0458	PE2	Control	Dnmt1	Dnmt1	+5 G to T		GCGGGCTGGAGCTGTTCGCGC (SEQ ID NO: 5272)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGGAGCTGTTCG (SEQ ID NO: 5273)	CGCTTGCATCTTGCA (SEQ ID NO: 5274)	TGCAAGATGCAAGCGCGAACAGCTCCAG (SEQ ID NO: 5275)	0	0	0	0
49E	N2A	Chen2021_0459	PE2	hMLH1 Δ754-756 GS codon (human	Dnmt1	Dnmt1	+5 G to T		GCGGGCTGGAGCTGTTCGCGC (SEQ ID NO: 5272)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGGAGCTGTTCG (SEQ ID NO: 5273)	CGCTTGCATCTTGCA (SEQ ID NO: 5274)	TGCAAGATGCAAGCGCGAACAGCTCCAG (SEQ ID NO: 5275)	0	0	0	0
49E	N2A	Chen2021_0460	PE2	mMLH1 1-757 GS codon (mouse	Dnmt1	Dnmt1	+5 G to T		GCGGGCTGGAGCTGTTCGCGC (SEQ ID NO: 5272)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGGAGCTGTTCG (SEQ ID NO: 5273)	CGCTTGCATCTTGCA (SEQ ID NO: 5274)	TGCAAGATGCAAGCGCGAACAGCTCCAG (SEQ ID NO: 5275)	0	0	0	0
49E	N2A	Chen2021_0461	PE2	Control	Dnmt1	Dnmt1	+5 G to T	+53	GCGGGCTGGAGCTGTTCGCGC (SEQ ID NO: 5272)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGGAGCTGTTCG (SEQ ID NO: 5273)	CGCTTGCATCTTGCA (SEQ ID NO: 5274)	TGCAAGATGCAAGCGCGAACAGCTCCAG (SEQ ID NO: 5275)	0	0	GCCGCGCGCGCGAAAAAGCCG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49E	N2A	Chen2021_0462	PE2	hMLH1 Δ754-756 GS codon (human MLH1dn)	Dnmt1	Dnmt1	+5 G to T	+53	GCGGGCTGGAGCTGTTCGCGC (SEQ ID NO: 5272)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGGAGCTGTTCG (SEQ ID NO: 5273)	CGCTTGCATCTTGCA (SEQ ID NO: 5274)	TGCAAGATGCAAGCGCGAACAGCTCCAG (SEQ ID NO: 5275)	0	0	GCCGCGCGCGCGAAAAAGCCG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49E	N2A	Chen2021_0463	PE2	mMLH1 1-757 GS codon (mouse MLH1dn)	Dnmt1	Dnmt1	+5 G to T	+53	GCGGGCTGGAGCTGTTCGCGC (SEQ ID NO: 5272)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGGAGCTGTTCG (SEQ ID NO: 5273)	CGCTTGCATCTTGCA (SEQ ID NO: 5274)	TGCAAGATGCAAGCGCGAACAGCTCCAG (SEQ ID NO: 5275)	0	0	GCCGCGCGCGCGAAAAAGCCG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49E	N2A	Chen2021_0464	PE2	Control	Chd2	Chd2	+5 G to A		GCGGTAGCTCCCAGAACGGT (SEQ ID NO: 5277)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TAGCTCCCAGAAC (SEQ ID NO: 5278)	GGTGAGCATC (SEQ ID NO: 5279)	GATGCTCACCGTTCTGGGAGCTA (SEQ ID NO: 5280)	0	0	0	0
49E	N2A	Chen2021_0465	PE2	hMLH1 Δ754-756 GS codon (human	Chd2	Chd2	+5 G to A		GCGGTAGCTCCCAGAACGGT (SEQ ID NO: 5277)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TAGCTCCCAGAAC (SEQ ID NO: 5278)	GGTGAGCATC (SEQ ID NO: 5279)	GATGCTCACCGTTCTGGGAGCTA (SEQ ID NO: 5280)	0	0	0	0
49E	N2A	Chen2021_0466	PE2	mMLH1 1-757 GS codon (mouse	Chd2	Chd2	+5 G to A		GCGGTAGCTCCCAGAACGGT (SEQ ID NO: 5277)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TAGCTCCCAGAAC (SEQ ID NO: 5278)	GGTGAGCATC (SEQ ID NO: 5279)	GATGCTCACCGTTCTGGGAGCTA (SEQ ID NO: 5280)	0	0	0	0
49E	N2A	Chen2021_0467	PE2	Control	Chd2	Chd2	+5 G to A	+52	GCGGTAGCTCCCAGAACGGT (SEQ ID NO: 5277)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TAGCTCCCAGAAC (SEQ ID NO: 5278)	GGTGAGCATC (SEQ ID NO: 5279)	GATGCTCACCGTTCTGGGAGCTA (SEQ ID NO: 5280)	0	0	GACCATCAGTATGAGCAGCAT (SEQ ID NO: 5281)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49E	N2A	Chen2021_0468	PE2	hMLH1 Δ754-756 GS codon (human MLH1dn)	Chd2	Chd2	+5 G to A	+52	GCGGTAGCTCCCAGAACGGT (SEQ ID NO: 5277)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TAGCTCCCAGAAC (SEQ ID NO: 5278)	GGTGAGCATC (SEQ ID NO: 5279)	GATGCTCACCGTTCTGGGAGCTA (SEQ ID NO: 5280)	0	0	GACCATCAGTATGAGCAGCAT (SEQ ID NO: 5281)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49E	N2A	Chen2021_0469	PE2	mMLH1 1-757 GS codon (mouse MLH1dn)	Chd2	Chd2	+5 G to A	+52	GCGGTAGCTCCCAGAACGGT (SEQ ID NO: 5277)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TAGCTCCCAGAAC (SEQ ID NO: 5278)	GGTGAGCATC (SEQ ID NO: 5279)	GATGCTCACCGTTCTGGGAGCTA (SEQ ID NO: 5280)	0	0	GACCATCAGTATGAGCAGCAT (SEQ ID NO: 5281)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49E	N2A	Chen2021_0470	PE2	Control	Col12a1	Col12a1	+2 A to C		GTGACTTCCATGGTTCCACAA (SEQ ID NO: 5282)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTCCATGGTTCCA (SEQ ID NO: 5283)	CCATGGGTCCATT (SEQ ID NO: 5284)	AATGGACCCATGGTGGAACCATGGAA (SEQ ID NO: 5285)	0	0	0	0

49E	N2A	Chen2021_0477	PE2	nML-H1 Δ754-756 GS codon (human	Cdl12a1	+2 A to C	-94	GTTGACTTCCAT GGTTCCACAA (SEQ ID NO: 5282)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTCCATGGTTCCA (SEQ ID NO: 5283)	CCATGGGTCCAATT (SEQ ID NO: 5284)	AATGGACCATGGTGGA (SEQ ID NO: 5285)	0	0	0	0	0	0	
49E	N2A	Chen2021_0472	PE2	nML-H1 1-757 GS codon (mouse	Cdl12a1	+2 A to C	-94	GTTGACTTCCAT GGTTCCACAA (SEQ ID NO: 5282)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTCCATGGTTCCA (SEQ ID NO: 5283)	CCATGGGTCCAATT (SEQ ID NO: 5284)	AATGGACCATGGTGGA (SEQ ID NO: 5285)	0	0	0	0	0	0	
49E	N2A	Chen2021_0473	PE2	Control	Cdl12a1	+2 A to C	-94	GTTGACTTCCAT GGTTCCACAA (SEQ ID NO: 5282)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTCCATGGTTCCA (SEQ ID NO: 5283)	CCATGGGTCCAATT (SEQ ID NO: 5284)	AATGGACCATGGTGGA (SEQ ID NO: 5285)	0	0	0	GCCTGAGCA (SEQ ID NO: 5286)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	
49E	N2A	Chen2021_0474	PE2	nML-H1 Δ754-756 GS codon (human	Cdl12a1	+2 A to C	-94	GTTGACTTCCAT GGTTCCACAA (SEQ ID NO: 5282)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTCCATGGTTCCA (SEQ ID NO: 5283)	CCATGGGTCCAATT (SEQ ID NO: 5284)	AATGGACCATGGTGGA (SEQ ID NO: 5285)	0	0	0	GCCTGAGCA (SEQ ID NO: 5286)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	
49E	N2A	Chen2021_0475	PE2	nML-H1 1-757 GS codon (mouse	Cdl12a1	+2 A to C	-94	GTTGACTTCCAT GGTTCCACAA (SEQ ID NO: 5282)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTCCATGGTTCCA (SEQ ID NO: 5283)	CCATGGGTCCAATT (SEQ ID NO: 5284)	AATGGACCATGGTGGA (SEQ ID NO: 5285)	0	0	0	GCCTGAGCA (SEQ ID NO: 5286)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	
49E	N2A	Chen2021_0476	PE2	Control	Cbh1b1	+6 G to A		GAGGGTTGCCC TTGCCACTCA (SEQ ID NO: 5287)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCCTTGCCAC (SEQ ID NO: 5288)	TCAGGAAAGGAGC (SEQ ID NO: 5289)	CCTCCTTTCCTGAGTGG (SEQ ID NO: 5290)	0	0	0	0	0	0	
49E	N2A	Chen2021_0477	PE2	nML-H1 Δ754-756 GS codon (human	Cbh1b1	+6 G to A		GAGGGTTGCCC TTGCCACTCA (SEQ ID NO: 5287)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCCTTGCCAC (SEQ ID NO: 5288)	TCAGGAAAGGAGC (SEQ ID NO: 5289)	CCTCCTTTCCTGAGTGG (SEQ ID NO: 5290)	0	0	0	0	0	0	
49E	N2A	Chen2021_0478	PE2	nML-H1 1-757 GS codon (mouse	Cbh1b1	+6 G to A		GAGGGTTGCCC TTGCCACTCA (SEQ ID NO: 5287)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCCTTGCCAC (SEQ ID NO: 5288)	TCAGGAAAGGAGC (SEQ ID NO: 5289)	CCTCCTTTCCTGAGTGG (SEQ ID NO: 5290)	0	0	0	0	0	0	
49E	N2A	Chen2021_0479	PE2	Control	Cbh1b1	+6 G to A	+70	GAGGGTTGCCC TTGCCACTCA (SEQ ID NO: 5287)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCCTTGCCAC (SEQ ID NO: 5288)	TCAGGAAAGGAGC (SEQ ID NO: 5289)	CCTCCTTTCCTGAGTGG (SEQ ID NO: 5290)	0	0	0	GAAAAAGCTG (SEQ ID NO: 5291)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	
49E	N2A	Chen2021_0480	PE2	nML-H1 Δ754-756 GS codon (human	Cbh1b1	+6 G to A	+70	GAGGGTTGCCC TTGCCACTCA (SEQ ID NO: 5287)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCCTTGCCAC (SEQ ID NO: 5288)	TCAGGAAAGGAGC (SEQ ID NO: 5289)	CCTCCTTTCCTGAGTGG (SEQ ID NO: 5290)	0	0	0	GAAAAAGCTG (SEQ ID NO: 5291)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	
49E	N2A	Chen2021_0481	PE2	nML-H1 1-757 GS codon (mouse	Cbh1b1	+6 G to A	+70	GAGGGTTGCCC TTGCCACTCA (SEQ ID NO: 5287)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCCTTGCCAC (SEQ ID NO: 5288)	TCAGGAAAGGAGC (SEQ ID NO: 5289)	CCTCCTTTCCTGAGTGG (SEQ ID NO: 5290)	0	0	0	GAAAAAGCTG (SEQ ID NO: 5291)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	
49F	HEK293T	Chen2021_0482	PE2	FANCF	FANCF	+7 A to C		GGAATCCGCTTC TCGAGCACC (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5296)	AGCTTGGCTGCTTT TCC (SEQ ID NO: 5292)	GGAAAAAGCGAGCCAGGT (SEQ ID NO: 5293)	0	0	0	0	0	0	
49F	HEK293T	Chen2021_0483	PE4	FANCF	FANCF	+7 A to C		GGAATCCGCTTC TCGAGCACC (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5296)	AGCTTGGCTGCTTT TCC (SEQ ID NO: 5292)	GGAAAAAGCGAGCCAGGT (SEQ ID NO: 5293)	0	0	0	0	0	0	
49F	HEK293T	Chen2021_0484	PE3	FANCF	FANCF	+7 A to C	-48	GGAATCCGCTTC TCGAGCACC (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5296)	AGCTTGGCTGCTTT TCC (SEQ ID NO: 5292)	GGAAAAAGCGAGCCAGGT (SEQ ID NO: 5293)	0	GGGGTGCCGA (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG GCTGCTAACTTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	
49F	HEK293T	Chen2021_0485	PE5	FANCF	FANCF	+7 A to C	-48	GGAATCCGCTTC TCGAGCACC (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5296)	AGCTTGGCTGCTTT TCC (SEQ ID NO: 5292)	GGAAAAAGCGAGCCAGGT (SEQ ID NO: 5293)	0	GGGGTGCCGA (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG GGTGCTAACTTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	

49F	HEK293T	Chen2021_0486	PE2	PE3b	FANCF	FANCF	+7 A to C	+7	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGCTCGCTTTTCC (SEQ ID NO: 5292)	GGAAAAGCGAGCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5293)	0	0	GAAGCTCGGAAAAGCGAGCC (SEQ ID NO: 5294)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49F	HEK293T	Chen2021_0487	PE2	PE5b	FANCF	FANCF	+7 A to C	+7	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGCTCGCTTTTCC (SEQ ID NO: 5292)	GGAAAAGCGAGCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5293)	0	0	GAAGCTCGGAAAAGCGAGCC (SEQ ID NO: 5294)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0488	PE2	PE2	HEK4	HEK4	+2 G to T		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	0	0
49G	HEK293T	Chen2021_0489	PE2	PE3	HEK4	HEK4	+2 G to T	-95	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GTCCCTTCCTTCCACCCAGCC (SEQ ID NO: 5299)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0490	PE2	PE3	HEK4	HEK4	+2 G to T	-52	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GCCCTGCCTGTCATCCTGCTT (SEQ ID NO: 5300)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0491	PE2	PE3	HEK4	HEK4	+2 G to T	-26	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GCAGTGCCACCGGGGCGCCG (SEQ ID NO: 5301)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0492	PE2	PE3	HEK4	HEK4	+2 G to T	+52	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GCGGGGGCTCAGAGAGGGCA (SEQ ID NO: 5302)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0493	PE2	PE3	HEK4	HEK4	+2 G to T	+74	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GACACACACACAGGCCTGG (SEQ ID NO: 5303)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0494	PE2	PE4	HEK4	HEK4	+2 G to T		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	0	0
49G	HEK293T	Chen2021_0495	PE2	PE5	HEK4	HEK4	+2 G to T	-95	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GTCCCTTCCTTCCACCCAGCC (SEQ ID NO: 5299)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0496	PE2	PE5	HEK4	HEK4	+2 G to T	-52	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GCCCTGCCTGTCATCCTGCTT (SEQ ID NO: 5300)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0497	PE2	PE5	HEK4	HEK4	+2 G to T	-26	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GCAGTGCCACCGGGGCGCCG (SEQ ID NO: 5301)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0498	PE2	PE5	HEK4	HEK4	+2 G to T	+52	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GCGGGGGCTCAGAGAGGGCA (SEQ ID NO: 5302)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
49G	HEK293T	Chen2021_0499	PE2	PE5	HEK4	HEK4	+2 G to T	+74	GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TTGGGGTTAA (SEQ ID NO: 5297)	TTAACCCCAACCTCCAGCC (SEQ ID NO: 5298)	0	0	GACACACACACAGGCCTGG (SEQ ID NO: 5303)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0500	PE2	PE2	FANCF	FANCF	+1 A to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	GCCTGGATCGCTTTTCC (SEQ ID NO: 5304)	GGAAAAGCGATCCAGGCGCTGCAGAAGGGAT (SEQ ID NO: 5305)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0501	PE2	PE2	FANCF	FANCF	+1 A to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	TCCTGGATCGCTTTTCC (SEQ ID NO: 5306)	GGAAAAGCGATCCAGGAGCTGCAGAAGGGAT (SEQ ID NO: 5307)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0502	PE2	PE2	FANCF	FANCF	+2 C to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	AACTGGATCGCTTTTCC (SEQ ID NO: 5308)	GGAAAAGCGATCCAGTTGCTGCAGAAGGGAT (SEQ ID NO: 5309)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0503	PE2	PE2	FANCF	FANCF	+3 C to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGTGGATCGCTTTTCC (SEQ ID NO: 5310)	GGAAAAGCGATCCACGTGCTGCAGAAGGGAT (SEQ ID NO: 5311)	0	0	0	0
44A,44B,44G,444G,50A,50B,50G,51D	HEK293T	Chen2021_0504	PE2	PE2	FANCF	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0505	PE2	PE2	FANCF	FANCF	+4 T to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCAGGATCGCTTTTCC (SEQ ID NO: 5257)	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT (SEQ ID NO: 5258)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0506	PE2	PE2	FANCF	FANCF	+4 T to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGGGATCGCTTTTCC (SEQ ID NO: 5314)	GGAAAAGCGATCCCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5315)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0507	PE2	PE2	FANCF	FANCF	+5 G to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTAGATCGCTTTTCC (SEQ ID NO: 5316)	GGAAAAGCGATCTAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5317)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0508	PE2	PE2	FANCF	FANCF	+6 G to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGCATCGCTTTTCC (SEQ ID NO: 5318)	GGAAAAGCGATGCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5319)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0509	PE2	PE2	FANCF	FANCF	+7 A to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGCTCGCTTTTCC (SEQ ID NO: 5292)	GGAAAAGCGAGCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5293)	0	0	0	0
44A,44B,50A,50B,50G,51D,51E	HEK293T	Chen2021_0510	PE2	PE2	FANCF	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0511	PE2	PE2	FANCF	FANCF	+10 G to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATCTCTTTTCC (SEQ ID NO: 5322)	GGAAAAGAGATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5323)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0512	PE2	PE4	FANCF	FANCF	+1 A to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	GCCTGGATCGCTTTTCC (SEQ ID NO: 5304)	GGAAAAGCGATCCAGGCGCTGCAGAAGGGAT (SEQ ID NO: 5305)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0513	PE2	PE4	FANCF	FANCF	+1 A to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	TCCTGGATCGCTTTTCC (SEQ ID NO: 5306)	GGAAAAGCGATCCAGGAGCTGCAGAAGGGAT (SEQ ID NO: 5307)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0514	PE2	PE4	FANCF	FANCF	+2 C to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	AACTGGATCGCTTTTCC (SEQ ID NO: 5308)	GGAAAAGCGATCCAGTTGCTGCAGAAGGGAT (SEQ ID NO: 5309)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0515	PE2	PE4	FANCF	FANCF	+3 C to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGTGGATCGCTTTTCC (SEQ ID NO: 5310)	GGAAAAGCGATCCACGTGCTGCAGAAGGGAT (SEQ ID NO: 5311)	0	0	0	0
44A,44B,44G,444G,50A,50B,50G,51D	HEK293T	Chen2021_0516	PE2	PE4	FANCF	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0

44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0517	PE2	PE4	FANCF	FANCF	+4 T to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCAGGATCGCTTTTCC (SEQ ID NO: 5257)	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT (SEQ ID NO: 5258)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0518	PE2	PE4	FANCF	FANCF	+4 T to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGGGATCGCTTTTCC (SEQ ID NO: 5314)	GGAAAAGCGATCCCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5315)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0519	PE2	PE4	FANCF	FANCF	+5 G to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTAGATCGCTTTTCC (SEQ ID NO: 5316)	GGAAAAGCGATCTAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5317)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0520	PE2	PE4	FANCF	FANCF	+6 G to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGCATCGCTTTTCC (SEQ ID NO: 5318)	GGAAAAGCGATGCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5319)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0521	PE2	PE4	FANCF	FANCF	+7 A to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGCTCGCTTTTCC (SEQ ID NO: 5292)	GGAAAAGCGAGCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5293)	0	0	0	0
44A,44B,50A,50B,50G,51D,51E	HEK293T	Chen2021_0522	PE2	PE4	FANCF	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0523	PE2	PE4	FANCF	FANCF	+10 G to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATCTCTTTTCC (SEQ ID NO: 5322)	GGAAAAGAGATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5323)	0	0	0	0
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0524	PE2	PE3	FANCF	FANCF	+1 A to G	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	GCCTGGATCGCTTTTCC (SEQ ID NO: 5304)	GGAAAAGCGATCCAGGCGCTGCAGAAGGGAT (SEQ ID NO: 5305)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0525	PE2	PE3	FANCF	FANCF	+1 A to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	TCCTGGATCGCTTTTCC (SEQ ID NO: 5306)	GGAAAAGCGATCCAGGAGCTGCAGAAGGGAT (SEQ ID NO: 5307)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0526	PE2	PE3	FANCF	FANCF	+2 C to A	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	AACTGGATCGCTTTTCC (SEQ ID NO: 5308)	GGAAAAGCGATCCAGTTGCTGCAGAAGGGAT (SEQ ID NO: 5309)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,50A,50B,50G,51B,51D	HEK293T	Chen2021_0527	PE2	PE3	FANCF	FANCF	+3 C to G	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGTGGATCGCTTTTCC (SEQ ID NO: 5310)	GGAAAAGCGATCCACGTGCTGCAGAAGGGAT (SEQ ID NO: 5311)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,50A,50B,50G,51B,51D	HEK293T	Chen2021_0528	PE2	PE3	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0529	PE2	PE3	FANCF	FANCF	+4 T to A	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCAGGATCGCTTTTCC (SEQ ID NO: 5257)	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT (SEQ ID NO: 5258)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,50A,50B,50G,51B,51D	HEK293T	Chen2021_0530	PE2	PE3	FANCF	FANCF	+4 T to G	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGGGATCGCTTTTCC (SEQ ID NO: 5314)	GGAAAAGCGATCCCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5315)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50C,50G,51B,51D	HEK293T	Chen2021_0531	PE2	PE3	FANCF	FANCF	+5 G to A	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTAGATCGCTTTTCC (SEQ ID NO: 5316)	GGAAAAGCGATCTAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5317)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A,44B,50A,50B,50C,50G,51B,51D	HEK293T	Chen2021_0532	PE2	PE3	FANCF	FANCF	+6 G to C	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGCATCGCTTTTCC (SEQ ID NO: 5318)	GGAAAAGCGATGCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5319)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0533	PE2	PE3	FANCF	FANCF	+7 A to C	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGCTCGCTTTTCC (SEQ ID NO: 5292)	GGAAAAGCGAGCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5293)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0534	PE2	PE3	FANCF	FANCF	+8 T to C	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0535	PE2	PE3	FANCF	FANCF	+10 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATCTCTTTTCC (SEQ ID NO: 5322)	GGAAAAGAGATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5323)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0536	PE2	PE5	FANCF	FANCF	+1 A to G	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	GCCTGGATCGCTTTTCC (SEQ ID NO: 5304)	GGAAAAGCGATCCAGGCGCTGCAGAAGGGAT (SEQ ID NO: 5305)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0537	PE2	PE5	FANCF	FANCF	+1 A to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	TCCTGGATCGCTTTTCC (SEQ ID NO: 5306)	GGAAAAGCGATCCAGGAGCTGCAGAAGGGAT (SEQ ID NO: 5307)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0538	PE2	PE5	FANCF	FANCF	+2 C to A	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	AACTGGATCGCTTTTCC (SEQ ID NO: 5308)	GGAAAAGCGATCCAGTTGCTGCAGAAGGGAT (SEQ ID NO: 5309)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,50A,50B,50G,51B,51D	HEK293T	Chen2021_0539	PE2	PE5	FANCF	FANCF	+3 C to G	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGTGGATCGCTTTTCC (SEQ ID NO: 5310)	GGAAAAGCGATCCACGTGCTGCAGAAGGGAT (SEQ ID NO: 5311)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,50A,50B,50G,51B,51D	HEK293T	Chen2021_0540	PE2	PE5	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0541	PE2	PE5	FANCF	FANCF	+4 T to A	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCAGGATCGCTTTTCC (SEQ ID NO: 5257)	GGAAAAGCGATCCTGGTGCTGCAGAAGGGAT (SEQ ID NO: 5258)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,50A,50B,50G,51B,51D	HEK293T	Chen2021_0542	PE2	PE5	FANCF	FANCF	+4 T to G	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGGGATCGCTTTTCC (SEQ ID NO: 5314)	GGAAAAGCGATCCCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5315)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50C,50G,51B,51D	HEK293T	Chen2021_0543	PE2	PE5	FANCF	FANCF	+5 G to A	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTAGATCGCTTTTCC (SEQ ID NO: 5316)	GGAAAAGCGATCTAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5317)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50C,50G,51B,51D	HEK293T	Chen2021_0544	PE2	PE5	FANCF	FANCF	+6 G to C	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGCATCGCTTTTCC (SEQ ID NO: 5318)	GGAAAAGCGATGCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5319)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0545	PE2	PE5	FANCF	FANCF	+7 A to C	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGCTCGCTTTTCC (SEQ ID NO: 5292)	GGAAAAGCGAGCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5293)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0546	PE2	PE5	FANCF	FANCF	+8 T to C	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51B,51D	HEK293T	Chen2021_0547	PE2	PE5	FANCF	FANCF	+10 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATCTCTTTTCC (SEQ ID NO: 5322)	GGAAAAGAGATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5323)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,444G,50A,50B,50G,51D	HEK293T	Chen2021_0548	PE2	PE2	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0549	PE2	PE2	HEK3	HEK3	+1 T to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5324)	TCTGCCATCGCGTGCTCAGTCTG (SEQ ID NO: 5325)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0550	PE2	PE2	HEK3	HEK3	+1 T to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGATGGCAGA (SEQ ID NO: 5326)	TCTGCCATCCCGTGCTCAGTCTG (SEQ ID NO: 5327)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0551	PE2	PE2	HEK3	HEK3	+3 A to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGCTGGCAGA (SEQ ID NO: 5328)	TCTGCCAGCACGTGCTCAGTCTG (SEQ ID NO: 5329)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0552	PE2	PE2	HEK3	HEK3	+3 A to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGTGGCAGA (SEQ ID NO: 5330)	TCTGCCACCACGTGCTCAGTCTG (SEQ ID NO: 5331)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0553	PE2	PE2	HEK3	HEK3	+3 A to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5332)	TCTGCCAACACGTGCTCAGTCTG (SEQ ID NO: 5333)	0	0	0	0
44A,44B,44G,50A,50B,50C,50G,51D	HEK293T	Chen2021_0554	PE2	PE2	HEK3	HEK3	+5 G to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATAGCAGA (SEQ ID NO: 5334)	TCTGCTATCACGTGCTCAGTCTG (SEQ ID NO: 5335)	0	0	0	0
44A,44B,44G,50A,50B,50C,50G,51D	HEK293T	Chen2021_0555	PE2	PE2	HEK3	HEK3	+5 G to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGCAGA (SEQ ID NO: 5336)	TCTGCGATCACGTGCTCAGTCTG (SEQ ID NO: 5337)	0	0	0	0
44A,44B,44G,50A,50B,50C,50G,51D,51E	HEK293T	Chen2021_0556	PE2	PE2	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0557	PE2	PE2	HEK3	HEK3	+7 C to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGAAGA (SEQ ID NO: 5338)	TCTTCCATCACGTGCTCAGTCTG (SEQ ID NO: 5339)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0558	PE2	PE2	HEK3	HEK3	+7 C to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGGAGA (SEQ ID NO: 5340)	TCTCCCATCACGTGCTCAGTCTG (SEQ ID NO: 5341)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0559	PE2	PE2	HEK3	HEK3	+7 C to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTAGA (SEQ ID NO: 5342)	TCTACCATCACGTGCTCAGTCTG (SEQ ID NO: 5343)	0	0	0	0
44A,44B,44G,444G,50A,50B,50G,51D	HEK293T	Chen2021_0560	PE2	PE4	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0

44A 44B 44G 50A 50B 50G T 51D	HEK293 Chen2021_0561	Chen2021	PE2	PE4	HEK3	HEK3	+1 T to C	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5324)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5325)	0	0	0			
44A 44B 44G 50A 50B 50G T 51D	HEK293 Chen2021_0562	Chen2021	PE2	PE4	HEK3	HEK3	+1 T to G	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5326)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5327)	0	0	0			
44A 44B 44G 50A 50B 50G T 51D	HEK293 Chen2021_0563	Chen2021	PE2	PE4	HEK3	HEK3	+3 A to C	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5328)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5329)	0	0	0			
44A 44B 44G 50A 50B 50G T 51D	HEK293 Chen2021_0564	Chen2021	PE2	PE4	HEK3	HEK3	+3 A to G	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5330)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5331)	0	0	0			
44A 44B 44G 50A 50B 50G T 51D	HEK293 Chen2021_0565	Chen2021	PE2	PE4	HEK3	HEK3	+3 A to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5332)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5333)	0	0	0			
44A 44B 44G 50A 50B 50G T 51D	HEK293 Chen2021_0566	Chen2021	PE2	PE4	HEK3	HEK3	+5 G to A	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATAGCAGA (SEQ ID NO: 5334)	TCTGCTATCGACGTGCTCA GTCTG (SEQ ID NO: 5335)	0	0	0			
44A 44B 44G 50A 50B 50C 50G 51D	HEK293 Chen2021_0567	Chen2021	PE2	PE4	HEK3	HEK3	+5 G to C	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGCAGA (SEQ ID NO: 5336)	TCTGCGATCACGTGCTC AGTCTG (SEQ ID NO: 5337)	0	0	0			
44A 44B 44G 50A 50B 50C 50G 51D	HEK293 Chen2021_0568	Chen2021	PE2	PE4	HEK3	HEK3	+5 G to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCA GTCTG (SEQ ID NO: 5234)	0	0	0			
44A 44B 44G 50A 50B 50G 51D	HEK293 Chen2021_0569	Chen2021	PE2	PE4	HEK3	HEK3	+7 C to A	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGAAGA (SEQ ID NO: 5338)	TCTTCGATCACGTGCTCA GTCTG (SEQ ID NO: 5339)	0	0	0			
44A 44B 44G 50A 50B 50G T 51D	HEK293 Chen2021_0570	Chen2021	PE2	PE4	HEK3	HEK3	+7 C to G	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGGAGA (SEQ ID NO: 5340)	TCTCCGATCACGTGCTCA GTCTG (SEQ ID NO: 5341)	0	0	0			
44A 44B 44G 50A 50B 50G T 51D	HEK293 Chen2021_0571	Chen2021	PE2	PE4	HEK3	HEK3	+7 C to T	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGTAGA (SEQ ID NO: 5342)	TCTACGATCACGTGCTCA GTCTG (SEQ ID NO: 5343)	0	0	0			
44A 44B 44G 46E 46F 50A 50B 50G 51D	HEK293 Chen2021_0572	Chen2021	PE2	PE3	HEK3	HEK3	+1 T to A	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCGTCGTGCTCA GTCTG (SEQ ID NO: 5248)	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 44G 46E 46F 50A 50B 50G 51D	HEK293 Chen2021_0573	Chen2021	PE2	PE3	HEK3	HEK3	+1 T to C	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5324)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5325)	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 44G 46E 46F 50A 50B 50G 51D	HEK293 Chen2021_0574	Chen2021	PE2	PE3	HEK3	HEK3	+1 T to G	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGATGGCAGA (SEQ ID NO: 5326)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5327)	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 44G 46E 46F 50A 50B 50G 51D	HEK293 Chen2021_0575	Chen2021	PE2	PE3	HEK3	HEK3	+3 A to C	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGCTGGCAGA (SEQ ID NO: 5328)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5329)	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 44G 46E 46F 50A 50B 50G 51D	HEK293 Chen2021_0576	Chen2021	PE2	PE3	HEK3	HEK3	+3 A to G	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGTGGCAGA (SEQ ID NO: 5330)	TCTGCCATCGTCGTGCTC AGTCTG (SEQ ID NO: 5331)	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		

44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0577	PE2	PE3	HEK3	+3 A to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5332)	TCTGCCAAGACGTGCTCAG (SEQ ID NO: 5333)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0578	PE2	PE3	HEK3	+5 G to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATAGCAGA (SEQ ID NO: 5334)	TCTGCTATCAGTGCTCA (SEQ ID NO: 5335)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0579	PE2	PE3	HEK3	+5 G to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATCGCAGA (SEQ ID NO: 5336)	TCTGCGATCAGTGCTCA (SEQ ID NO: 5337)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0580	PE2	PE3	HEK3	+5 G to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCAGTGCTCA (SEQ ID NO: 5234)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0581	PE2	PE3	HEK3	+7 C to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATGGAAGA (SEQ ID NO: 5338)	TCTTCCATCAGGTGCTCA (SEQ ID NO: 5339)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0582	PE2	PE3	HEK3	+7 C to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATGGGAGA (SEQ ID NO: 5340)	TCTCCCATCAGGTGCTCA (SEQ ID NO: 5341)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0583	PE2	PE3	HEK3	+7 C to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATGGTAGA (SEQ ID NO: 5342)	TCTACCATCAGGTGCTCA (SEQ ID NO: 5343)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0584	PE2	PE5	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCA (SEQ ID NO: 5248)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0585	PE2	PE5	HEK3	+1 T to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5324)	TCTGCCATCGGTGCTCA (SEQ ID NO: 5325)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0586	PE2	PE5	HEK3	+1 T to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	GGATGGCAGA (SEQ ID NO: 5326)	TCTGCCATCCCGTGCTCA (SEQ ID NO: 5327)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0587	PE2	PE5	HEK3	+3 A to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGCTGGCAGA (SEQ ID NO: 5328)	TCTGCCAGCACGTGCTC (SEQ ID NO: 5329)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0588	PE2	PE5	HEK3	+3 A to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGGTGGCAGA (SEQ ID NO: 5330)	TCTGCCAGCACGTGCTC (SEQ ID NO: 5331)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293_0589	PE2	PE5	HEK3	+3 A to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5332)	TCTGCCAACACGTGCTCAG (SEQ ID NO: 5333)	10		GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A,44B,44G, 45E,45F,50A, 50B,50C,50G, 51B,51D	HEK293T	Chen2021_0590	PE2	PE5	HEK3	HEK3	+5 G to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATAGCAGA (SEQ ID NO: 5334)	TCTGCTATCACGTGCTCAGTCTG (SEQ ID NO: 5335)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G, 45E,45F,50A, 50B,50C,50G, 51B,51D	HEK293T	Chen2021_0591	PE2	PE5	HEK3	HEK3	+5 G to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGCAGA (SEQ ID NO: 5336)	TCTGCGATCACGTGCTCAGTCTG (SEQ ID NO: 5337)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G, 45E,45F,50A, 50B,50C,50G, 51B,51D	HEK293T	Chen2021_0592	PE2	PE5	HEK3	HEK3	+5 G to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G, 45E,45F,50A, 50B,50G,51B, 51D	HEK293T	Chen2021_0593	PE2	PE5	HEK3	HEK3	+7 C to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGAAGA (SEQ ID NO: 5338)	TCTTCCATCACGTGCTCAGTCTG (SEQ ID NO: 5339)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G, 45E,45F,50A, 50B,50G,51B, 51D	HEK293T	Chen2021_0594	PE2	PE5	HEK3	HEK3	+7 C to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGGAGA (SEQ ID NO: 5340)	TCTCCCATCACGTGCTCAGTCTG (SEQ ID NO: 5341)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G, 45E,45F,50A, 50B,50G,51B, 51D	HEK293T	Chen2021_0595	PE2	PE5	HEK3	HEK3	+7 C to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTAGA (SEQ ID NO: 5342)	TCTACCATCACGTGCTCAGTCTG (SEQ ID NO: 5343)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,444 G,50A,50B,50 G,51D	HEK293T	Chen2021_0596	PE2	PE2	EMX1	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0597	PE2	PE2	EMX1	EMX1	+2 A to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GTAGGGCTCCCATCAC (SEQ ID NO: 5346)	GTGATGGGAGCCCTACTTCTTCTGCTCGGA (SEQ ID NO: 5347)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0598	PE2	PE2	EMX1	EMX1	+3 A to G		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAGGGGCTCCCATCAC (SEQ ID NO: 5348)	GTGATGGGAGCCCCTCTTCTTCTGCTCGGA (SEQ ID NO: 5349)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0599	PE2	PE2	EMX1	EMX1	+4 G to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAACGGCTCCCATCAC (SEQ ID NO: 5350)	GTGATGGGAGCCGTTCTTCTTCTGCTCGGA (SEQ ID NO: 5351)	0	0	0	0
44A,44B,50A, 50B,50C,50G, 51D	HEK293T	Chen2021_0600	PE2	PE2	EMX1	EMX1	+5 G to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	0	0
44A,44B,50A, 50B,50C,50G, 51D,51E	HEK293T	Chen2021_0601	PE2	PE2	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0602	PE2	PE2	EMX1	EMX1	+7 C to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGATCCCATCAC (SEQ ID NO: 5352)	GTGATGGGATCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5353)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0603	PE2	PE2	EMX1	EMX1	+8 T to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCACCCATCAC (SEQ ID NO: 5354)	GTGATGGGTGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5355)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0604	PE2	PE2	EMX1	EMX1	+8 T to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCCCCCATCAC (SEQ ID NO: 5356)	GTGATGGGGGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5357)	0	0	0	0

44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0605	PE2	PE2	EMX1	EMX1	+8 T to G		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCGCCCATCAC (SEQ ID NO: 5358)	GTGATGGGCGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5359)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0606	PE2	PE2	EMX1	EMX1	+9 C to G		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCTGCCATCAC (SEQ ID NO: 5360)	GTGATGGCAGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5361)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0607	PE2	PE2	EMX1	EMX1	+9 C to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCTTCCATCAC (SEQ ID NO: 5362)	GTGATGGAAGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5363)	0	0	0	0
44A,44B,444 G,50A,50B,50 G,51D	HEK293T	Chen2021_0608	PE2	PE4	EMX1	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0609	PE2	PE4	EMX1	EMX1	+2 A to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GTAGGGCTCCCATCAC (SEQ ID NO: 5346)	GTGATGGGAGCCCTACTTCTTCTGCTCGGA (SEQ ID NO: 5347)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0610	PE2	PE4	EMX1	EMX1	+3 A to G		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAGGGGCTCCCATCAC (SEQ ID NO: 5348)	GTGATGGGAGCCCCTCTTCTTCTGCTCGGA (SEQ ID NO: 5349)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0611	PE2	PE4	EMX1	EMX1	+4 G to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAACGGCTCCCATCAC (SEQ ID NO: 5350)	GTGATGGGAGCCGTTCTTCTTCTGCTCGGA (SEQ ID NO: 5351)	0	0	0	0
44A,44B,50A, 50B,50C,50G, 51D	HEK293T	Chen2021_0612	PE2	PE4	EMX1	EMX1	+5 G to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	0	0
44A,44B,50A, 50B,50C,50G, 51D,51E	HEK293T	Chen2021_0613	PE2	PE4	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0614	PE2	PE4	EMX1	EMX1	+7 C to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGATCCCATCAC (SEQ ID NO: 5352)	GTGATGGGATCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5353)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0615	PE2	PE4	EMX1	EMX1	+8 T to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCACCCATCAC (SEQ ID NO: 5354)	GTGATGGGTGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5355)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0616	PE2	PE4	EMX1	EMX1	+8 T to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCCCCCATCAC (SEQ ID NO: 5356)	GTGATGGGGGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5357)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0617	PE2	PE4	EMX1	EMX1	+8 T to G		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCGCCCATCAC (SEQ ID NO: 5358)	GTGATGGGCGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5359)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0618	PE2	PE4	EMX1	EMX1	+9 C to G		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCTGCCATCAC (SEQ ID NO: 5360)	GTGATGGCAGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5361)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0619	PE2	PE4	EMX1	EMX1	+9 C to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCTTCCATCAC (SEQ ID NO: 5362)	GTGATGGAAGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5363)	0	0	0	0
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293T	Chen2021_0620	PE2	PE3	EMX1	EMX1	+2 A to C	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0621	PE2	PE3	EMX1	+2 A to T	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GTACGGCTCCCATC (SEQ ID NO: 5346)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0622	PE2	PE3	EMX1	+3 A to G	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAGGGGCTCCCATC (SEQ ID NO: 5348)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0623	PE2	PE3	EMX1	+4 G to C	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAACGGCTCCCATC (SEQ ID NO: 5350)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0624	PE2	PE3	EMX1	+5 G to A	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGAGGCTCCCATC (SEQ ID NO: 5249)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0625	PE2	PE3	EMX1	+5 G to T	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGGCTCCCATC (SEQ ID NO: 5236)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0626	PE2	PE3	EMX1	+7 C to A	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGCGATCCCATC (SEQ ID NO: 5352)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0627	PE2	PE3	EMX1	+8 T to A	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGGGGCACCCATC (SEQ ID NO: 5354)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0628	PE2	PE3	EMX1	+8 T to C	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGGGGCCCCCATC (SEQ ID NO: 5355)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0629	PE2	PE3	EMX1	+8 T to G	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGGGGCGCCCAT (SEQ ID NO: CAC (SEQ ID NO: 5356)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0630	PE2	PE3	EMX1	+9 C to G	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGGGGCTGCCATC (SEQ ID NO: 5360)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0631	PE2	PE3	EMX1	+9 C to T	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGGGCTTCCCATC (SEQ ID NO: 5362)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0632	PE2	PE5	EMX1	+2 A to C	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GCAGGGCGTCCCATC (SEQ ID NO: AC (SEQ ID NO: 5344)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)
44A, 44B, 45E, 45F, 50A, 50B, 50G, 51B, 51D	HEK293T	Chen2021_0633	PE2	PE5	EMX1	+2 A to T	57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GTAGGGCTCCCATC (SEQ ID NO: AC (SEQ ID NO: 5346)	16		GCCGTTTGTCACTTTGTCCTCTGCTCGGGA (SEQ ID NO: 5241)

44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0634	PE2	PE5	EMX1	EMX1	+3 A to G	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAGGGGCTCCCATCAC (SEQ ID NO: 5348)	GTGATGGGAGCCCCTCTTCTTCTGCTCGGA (SEQ ID NO: 5349)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0635	PE2	PE5	EMX1	EMX1	+4 G to C	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAACGGCTCCCATCAC (SEQ ID NO: 5350)	GTGATGGGAGCCGTTCTTCTTCTGCTCGGA (SEQ ID NO: 5351)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50C,50G,51B,51D	HEK293T	Chen2021_0636	PE2	PE5	EMX1	EMX1	+5 G to A	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50C,50G,51B,51D	HEK293T	Chen2021_0637	PE2	PE5	EMX1	EMX1	+5 G to T	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0638	PE2	PE5	EMX1	EMX1	+7 C to A	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGATCCCATCAC (SEQ ID NO: 5352)	GTGATGGGATCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5353)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0639	PE2	PE5	EMX1	EMX1	+8 T to A	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCACCCATCAC (SEQ ID NO: 5354)	GTGATGGGTGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5355)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0640	PE2	PE5	EMX1	EMX1	+8 T to C	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCCCCCATCAC (SEQ ID NO: 5356)	GTGATGGGGGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5357)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0641	PE2	PE5	EMX1	EMX1	+8 T to G	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCGCCCATCAC (SEQ ID NO: 5358)	GTGATGGGCGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5359)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0642	PE2	PE5	EMX1	EMX1	+9 C to G	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCTGCCATCAC (SEQ ID NO: 5360)	GTGATGGCAGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5361)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0643	PE2	PE5	EMX1	EMX1	+9 C to T	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCTTCCATCAC (SEQ ID NO: 5362)	GTGATGGAAGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5363)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0644	PE2	PE2	RNF2	RNF2	+1 C to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5364)	AACGAACACCTCATGTAATGACTAAGATG (SEQ ID NO: 5365)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0645	PE2	PE2	RNF2	RNF2	+1 C to G		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	GTGAGGTGTTCGTT (SEQ ID NO. 5366)	AACGAACACCTCACGTAATGACTAAGATG (SEQ ID NO: 5367)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0646	PE2	PE2	RNF2	RNF2	+1 C to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TTGAGGTGTTCGTT (SEQ ID NO. 5253)	AACGAACACCTCAAGTAATGACTAAGATG (SEQ ID NO: 5254)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0647	PE2	PE2	RNF2	RNF2	+2 T to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CAGAGGTGTTCGTT (SEQ ID NO. 5368)	AACGAACACCTCTGGTAATGACTAAGATG (SEQ ID NO: 5369)	0	0	0	0

44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0648	PE2	PE2	RNF2	RNF2	+2 T to G		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CGGAGGTGTTCGTT (SEQ ID NO: 5370)	AACGAACACCTCCGGTAATGACTAAGATG (SEQ ID NO: 5371)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0649	PE2	PE2	RNF2	RNF2	+3 G to C		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTCAGGTGTTCGTT (SEQ ID NO: 5372)	AACGAACACCTGAGGTAATGACTAAGATG (SEQ ID NO: 5373)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0650	PE2	PE2	RNF2	RNF2	+4 A to C		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGCGGTGTTCGTT (SEQ ID NO: 5374)	AACGAACACCGCAGGTAATGACTAAGATG (SEQ ID NO: 5375)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0651	PE2	PE2	RNF2	RNF2	+4 A to G		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGGGGTGTTCGTT (SEQ ID NO: 5376)	AACGAACACCCCAGGTAATGACTAAGATG (SEQ ID NO: 5377)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0652	PE2	PE2	RNF2	RNF2	+4 A to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGTGGTGTTCGTT (SEQ ID NO: 5378)	AACGAACACCACAGGTAATGACTAAGATG (SEQ ID NO: 5379)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0653	PE2	PE2	RNF2	RNF2	+5 G to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGATGTGTTCGTT (SEQ ID NO: 5380)	AACGAACACATCAGGTAATGACTAAGATG (SEQ ID NO: 5381)	0	0	0	0
44A,44B,44G,50A,50B,50C,50G,51D	HEK293T	Chen2021_0654	PE2	PE2	RNF2	RNF2	+6 G to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0655	PE2	PE2	RNF2	RNF2	+7 T to C		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCGTTCGTT (SEQ ID NO: 5384)	AACGAACGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5385)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0656	PE2	PE4	RNF2	RNF2	+1 C to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5364)	AACGAACACCTCATGTAATGACTAAGATG (SEQ ID NO: 5365)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0657	PE2	PE4	RNF2	RNF2	+1 C to G		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	GTGAGGTGTTCGTT (SEQ ID NO: 5366)	AACGAACACCTCACGTAATGACTAAGATG (SEQ ID NO: 5367)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0658	PE2	PE4	RNF2	RNF2	+1 C to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TTGAGGTGTTCGTT (SEQ ID NO: 5253)	AACGAACACCTCAAGTAATGACTAAGATG (SEQ ID NO: 5254)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0659	PE2	PE4	RNF2	RNF2	+2 T to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CAGAGGTGTTCGTT (SEQ ID NO: 5368)	AACGAACACCTCTGGTAATGACTAAGATG (SEQ ID NO: 5369)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0660	PE2	PE4	RNF2	RNF2	+2 T to G		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CGGAGGTGTTCGTT (SEQ ID NO: 5370)	AACGAACACCTCCGGTAATGACTAAGATG (SEQ ID NO: 5371)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0661	PE2	PE4	RNF2	RNF2	+3 G to C		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTCAGGTGTTCGTT (SEQ ID NO: 5372)	AACGAACACCTGAGGTAATGACTAAGATG (SEQ ID NO: 5373)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0662	PE2	PE4	RNF2	RNF2	+4 A to C		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGCGGTGTTCGTT (SEQ ID NO: 5374)	AACGAACACCGCAGGTAATGACTAAGATG (SEQ ID NO: 5375)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0663	PE2	PE4	RNF2	RNF2	+4 A to G		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGGGGTGTTCGTT (SEQ ID NO: 5376)	AACGAACACCCCAGGTAATGACTAAGATG (SEQ ID NO: 5377)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0664	PE2	PE4	RNF2	RNF2	+4 A to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGTGGTGTTCGTT (SEQ ID NO: 5378)	AACGAACACCACAGGTAATGACTAAGATG (SEQ ID NO: 5379)	0	0	0	0

44A 44B 50A 50B 50C 50G 51D	HEK293T	Chen2021_0635	PE2	PE4	RNF2	+5 G to T	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTGATGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5380)	AAGGAACACATCAGGTTAAGTGACTAAGATG (SEQ ID NO: 5381)	10	0	0	0	
44A 44B 44G 50A 50B 50C 50G 51D	HEK293T	Chen2021_0636	PE2	PE4	RNF2	+6 G to A	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTGATGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5382)	AAGGAACACATCTGACGTTAAGTGACTAAGATG (SEQ ID NO: 5383)	10	0	0	0	
44A 44B 44G 50A 50B 50G 51D	HEK293T	Chen2021_0637	PE2	PE4	RNF2	+7 T to C	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTGATGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5384)	AAGGAACACCTCAGGTTAATGACTAAGATG (SEQ ID NO: 5385)	10	0	0	0	
44A 44B 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0638	PE3	PE2	RNF2	+1 C to A	GTCATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATGAGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5364)	AAGGAACACCTCATGTAAGTGACTAAGATG (SEQ ID NO: 5365)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0639	PE3	PE2	RNF2	+1 C to G	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTGATGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5386)	AAGGAACACCTCACGTTAAGTGACTAAGATG (SEQ ID NO: 5387)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0670	PE2	PE3	RNF2	+1 C to T	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	TTGAGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5253)	AAGGAACACCTCAAGTAAGTGACTAAGATG (SEQ ID NO: 5284)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0671	PE2	PE3	RNF2	+2 T to A	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	CACGAGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5388)	AAGGAACACCTCTGTGGTAAGTGACTAAGATG (SEQ ID NO: 5389)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0672	PE2	PE3	RNF2	+2 T to G	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	CGGAGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5370)	AAGGAACACCTCCGGTAATGACTAAGATG (SEQ ID NO: 5371)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0673	PE2	PE3	RNF2	+3 G to C	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTCGAGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5372)	AAGGAACACCTGAGGTAATGACTAAGATG (SEQ ID NO: 5373)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0674	PE2	PE3	RNF2	+4 A to C	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTCGCGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5374)	AAGGAACACCGCAGGTAATGACTAAGATG (SEQ ID NO: 5375)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0675	PE2	PE3	RNF2	+4 A to G	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTGGGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5376)	AAGGAACACCCCAGGTAATGACTAAGATG (SEQ ID NO: 5377)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 44G 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0676	PE2	PE3	RNF2	+4 A to T	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTGTGGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5378)	AAGGAACACCACAGGTAATGACTAAGATG (SEQ ID NO: 5379)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 45D 46E 46F 50A 50B 50C 50G 51D	HEK293T	Chen2021_0677	PE2	PE3	RNF2	+5 G to T	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	GTGATGTGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5380)	AAGGAACACATCTGAGGTTAAGTGACTAAGATG (SEQ ID NO: 5381)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A 44B 44G 45D 46E 46F 50A 50B 50G 51D	HEK293T	Chen2021_0678	PE2	PE3	RNF2	+6 G to A	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCAC (SEQ ID NO: 5292)	CTGAGATGTTCGTTTAAGACTAAGATG (SEQ ID NO: 5382)	AAGGAACATCTCAGGTTAAGTGACTAAGATG (SEQ ID NO: 5383)	10	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A,44B,44C,45D,45E,45F,50A,50B,50C,51D	HEK293T	Chen2021_0679	PE2	PE3	RNF2	+7 T to C	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTATA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CTGAGGGGTTTCGTT AAGCTAAGATG (SEQ ID NO: 5254)	AAGCGAAGCGCTCAGGTA ATCACTAAGATG (SEQ ID NO: 5386)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0680	PE2	PE5	RNF2	+1 C to A	-41	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	ATGAGGGTGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5254)	AAGCGAACGCCTCATGTAA TAAGCTAAGATG (SEQ ID NO: 5386)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0681	PE2	PE5	RNF2	+1 C to G	-41	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	GTGAGGGTGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5254)	AAGCGAACGCCTCAGGTAA TAAGCTAAGATG (SEQ ID NO: 5386)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0682	PE2	PE5	RNF2	+1 C to T	-41	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	TTGAGGGTGTTCGTT TGAGCTAAGATG (SEQ ID NO: 5254)	AAGCGAACGCCTCAAGTAA TGAGCTAAGATG (SEQ ID NO: 5386)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0683	PE2	PE5	RNF2	+2 T to A	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CAGAGGTGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5369)	AAGCGAACGCCTCTGGTAA TAAGCTAAGATG (SEQ ID NO: 5369)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0684	PE2	PE5	RNF2	+2 T to G	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CGGAGGTGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5370)	AAGCGAACGCCTCGGTAA ATGACTAAGATG (SEQ ID NO: 5371)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0685	PE2	PE5	RNF2	+3 G to C	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CTCAGGGTGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5372)	AAGCGAACGCCTGAGGTA ATGACTAAGATG (SEQ ID NO: 5373)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0686	PE2	PE5	RNF2	+4 A to C	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CTGGGGTGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5374)	AAGCGAACGCCGGAGGTA ATGACTAAGATG (SEQ ID NO: 5375)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0687	PE2	PE5	RNF2	+4 A to G	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CTGGGGTGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5376)	AAGCGAACGCCCGAGGTA ATGACTAAGATG (SEQ ID NO: 5377)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44C,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0688	PE2	PE5	RNF2	+4 A to T	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CTGTGGTGTTCGTT AAGCTAAGATG (SEQ ID NO: 5378)	AAGCGAACGCCACAAGGTA ATGACTAAGATG (SEQ ID NO: 5379)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50C,50G,51D	HEK293T	Chen2021_0689	PE2	PE5	RNF2	+5 G to T	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CTGATGTGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5380)	AAGCGAACGCATCAAGGTAA TGACTAAGATG (SEQ ID NO: 5381)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44C,45D,45E,45F,50A,50B,50C,50G,51D	HEK293T	Chen2021_0690	PE2	PE5	RNF2	+6 G to A	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CTGAGATGTTCGTT TAAGCTAAGATG (SEQ ID NO: 5382)	AAGCGAACATCTCAGGTAAA TGACTAAGATG (SEQ ID NO: 5383)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44C,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0691	PE2	PE5	RNF2	+7 T to C	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CATCTTAGTCATTA (SEQ ID NO: 5253)	CTGAGAGGCTTCGTT AAGCTAAGATG (SEQ ID NO: 5384)	AAGCGAACGCCTCAGGTAC ATCACTAAGATG (SEQ ID NO: 5385)	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0692	PE2	RUNX1	RUNX1	+1 C to A		CCATTTTCAGG AAGAAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAAGGA AG (SEQ ID NO: 5243)	AGATGGCGTTCACGAC A (SEQ ID NO: 5386)	TGTCTGAAGCCATCTCTTT ACTCTGAAAAAT (SEQ ID NO: 5387)	0	0	0

44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0693	PE2	PE2	RUNX1	RUNX1	+1 C to G		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	GGATGGCTTCAGACA (SEQ ID NO: 5388)	TGTCTGAAGCCATCCCTTCCTCCTGAAAAT (SEQ ID NO: 5389)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0694	PE2	PE2	RUNX1	RUNX1	+1 C to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	TGATGGCTTCAGACA (SEQ ID NO: 5390)	TGTCTGAAGCCATCACTTCCTCCTGAAAAT (SEQ ID NO: 5391)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0695	PE2	PE2	RUNX1	RUNX1	+2 G to A		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CAATGGCTTCAGACA (SEQ ID NO: 5392)	TGTCTGAAGCCATTGCTTCCTCCTGAAAAT (SEQ ID NO: 5393)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0696	PE2	PE2	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0697	PE2	PE2	RUNX1	RUNX1	+3 A to G		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGTGGCTTCAGACA (SEQ ID NO: 5394)	TGTCTGAAGCCACCGCTTCCTCCTGAAAAT (SEQ ID NO: 5395)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0698	PE2	PE2	RUNX1	RUNX1	+3 A to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGTTGGCTTCAGACA (SEQ ID NO: 5396)	TGTCTGAAGCCAACGCTTCCTCCTGAAAAT (SEQ ID NO: 5397)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0699	PE2	PE2	RUNX1	RUNX1	+4 T to A		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGAAGGCTTCAGACA (SEQ ID NO: 5398)	TGTCTGAAGCCTTCGCTTCCTCCTGAAAAT (SEQ ID NO: 5399)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0700	PE2	PE2	RUNX1	RUNX1	+4 T to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGACGGCTTCAGACA (SEQ ID NO: 5400)	TGTCTGAAGCCGTCGCTTCCTCCTGAAAAT (SEQ ID NO: 5401)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0701	PE2	PE2	RUNX1	RUNX1	+4 T to G		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGAGGGCTTCAGACA (SEQ ID NO: 5402)	TGTCTGAAGCCCTCGCTTCCTCCTGAAAAT (SEQ ID NO: 5403)	0	0	0	0
44A,44B,44G,50A,50B,50C,50G,51D	HEK293T	Chen2021_0702	PE2	PE2	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0703	PE2	PE2	RUNX1	RUNX1	+6 G to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATGCCTTCAGACA (SEQ ID NO: 5404)	TGTCTGAAGGCATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5405)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0704	PE2	PE4	RUNX1	RUNX1	+1 C to A		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	AGATGGCTTCAGACA (SEQ ID NO: 5386)	TGTCTGAAGCCATCTCTTCCTCCTGAAAAT (SEQ ID NO: 5387)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0705	PE2	PE4	RUNX1	RUNX1	+1 C to G		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	GGATGGCTTCAGACA (SEQ ID NO: 5388)	TGTCTGAAGCCATCCCTTCCTCCTGAAAAT (SEQ ID NO: 5389)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0706	PE2	PE4	RUNX1	RUNX1	+1 C to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	TGATGGCTTCAGACA (SEQ ID NO: 5390)	TGTCTGAAGCCATCACTTCCTCCTGAAAAT (SEQ ID NO: 5391)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0707	PE2	PE4	RUNX1	RUNX1	+2 G to A		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CAATGGCTTCAGACA (SEQ ID NO: 5392)	TGTCTGAAGCCATTGCTTCCTCCTGAAAAT (SEQ ID NO: 5393)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0708	PE2	PE4	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0709	PE2	PE4	RUNX1	RUNX1	+3 A to G		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGTGGCTTCAGACA (SEQ ID NO: 5394)	TGTCTGAAGCCACCGCTTCCTCCTGAAAAT (SEQ ID NO: 5395)	0	0	0	0

44A 44B 50A 50B 50G 51D	HEK293T	Chen2021_0710	PE2	PE4	RUNX1	RUNX1	+3 A to T	GCATTTCAGGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGTTGGCTTCAGAC A (SEQ ID NO: 5396)	TGTCTGAAGGCCAACGCTT CCTCGTCGAAAAT (SEQ ID NO: 5397)	0	0	0	0	0	0	0
44A 44B 44G 50A 50B 50G 51D	HEK293T	Chen2021_0711	PE2	PE4	RUNX1	RUNX1	+4 T to A	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGAAGGCTTCAGAC A (SEQ ID NO: 5398)	TGTCTGAAGGCCTTCGCTT CCTCGTCGAAAAT (SEQ ID NO: 5399)	0	0	0	0	0	0	0
44A 44B 50A 50B 50G 51D	HEK293T	Chen2021_0712	PE2	PE4	RUNX1	RUNX1	+4 T to C	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGAAGGCTTCAGAC A (SEQ ID NO: 5400)	TGTCTGAAGGCGTCGACT CCTCGTCGAAAAT (SEQ ID NO: 5401)	0	0	0	0	0	0	0
44A 44B 50A 50B 50G 51D	HEK293T	Chen2021_0713	PE2	PE4	RUNX1	RUNX1	+4 T to G	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGAGGGCTTCAGAC A (SEQ ID NO: 5402)	TGTCTGAAGGCCTCGCTT CCTCGTCGAAAAT (SEQ ID NO: 5403)	0	0	0	0	0	0	0
44A 44B 44G 50A 50B 50G 51D	HEK293T	Chen2021_0714	PE2	PE4	RUNX1	RUNX1	+5 G to T	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5359)	TGTCTGAAGCAATCGGCTT CCTCGTCGAAAAT (SEQ ID NO: 5239)	0	0	0	0	0	0	0
44A 44B 50A 50B 50G 51D	HEK293T	Chen2021_0715	PE2	PE4	RUNX1	RUNX1	+5 G to C	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGATTGCCTTCAGAC A (SEQ ID NO: 5404)	TGTCTGAAGGCATCGCTT CCTCGTCGAAAAT (SEQ ID NO: 5405)	0	0	0	0	0	0	0
44A 44B 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0716	PE3	RUNX1	RUNX1	+1 C to A	-38	GCATTTCAGGG AGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	AGATTGGCTTCAGAC A (SEQ ID NO: 5386)	TGTCTGAAGCCATCTCTTT CCTCGTCGAAAAT (SEQ ID NO: 5387)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0717	PE3	RUNX1	RUNX1	+1 C to G	-38	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGATGGCTTCAGAC A (SEQ ID NO: 5388)	TGTCTGAAGCCATCGCTT CCTCGTCGAAAAT (SEQ ID NO: 5389)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0718	PE3	RUNX1	RUNX1	+1 C to T	-38	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	TGATTGGCTTCAGAC A (SEQ ID NO: 5390)	TGTCTGAAGCCATCAGTT CCTCGTCGAAAAT (SEQ ID NO: 5391)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0719	PE3	RUNX1	RUNX1	+2 G to A	-38	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGCGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CAATGGCTTCAGAC A (SEQ ID NO: 5392)	TGTCTGAAGCCATTGCTT CCTCGTCGAAAAT (SEQ ID NO: 5393)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 44G 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0720	PE3	RUNX1	RUNX1	+3 A to C	-38	CCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGCGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGCTGTGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCACGCGCT CCTCGTCGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0721	PE3	RUNX1	RUNX1	+3 A to G	-38	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGCGTGGCTTCAGAC A (SEQ ID NO: 5394)	TGTCTGAAGCGCACCGGCT CCTCGTCGAAAAT (SEQ ID NO: 5395)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0722	PE2	PE3	RUNX1	RUNX1	+3 A to T	-38	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGTTGGCTTCAGAC A (SEQ ID NO: 5396)	TGTCTGAAGGCGTCGACT CCTCGTCGAAAAT (SEQ ID NO: 5397)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 44G 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0723	PE2	PE3	RUNX1	RUNX1	+4 T to A	-38	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGAAGGCTTCAGAC A (SEQ ID NO: 5398)	TGTCTGAAGGCCTTCGCTT CCTCGTCGAAAAT (SEQ ID NO: 5399)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		
44A 44B 45E 45F 50A 50B 50G 51B 51D	HEK293T	Chen2021_0724	PE3	RUNX1	RUNX1	+4 T to C	-38	GCATTTCAGGG AAGGAAGCGA (SEQ ID NO: 5240)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGCGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGAGGAGA AG (SEQ ID NO: 5243)	CGACGGCTTCAGAC A (SEQ ID NO: 5400)	TGTCTGAAGCGTCGACT CCTCGTCGAAAAT (SEQ ID NO: 5401)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)		

44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0725	PE2	PE3	RUNX1	RUNX1	+4 T to G	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGAGGGCTTCAGAC A (SEQ ID NO: 5402)	TGTCTGAAGCCCTCGCTT CCTCCTGAAAAT (SEQ ID NO: 5403)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G, 45E,45F,50A, 50B,50C,50G, 51B,51D	HEK293 T	Chen2021 _0726	PE2	PE3	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50C,50G,51B, 51D	HEK293 T	Chen2021 _0727	PE2	PE3	RUNX1	RUNX1	+6 G to C	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATGCCTTCAGAC A (SEQ ID NO: 5404)	TGTCTGAAGGCATCGCTT CCTCCTGAAAAT (SEQ ID NO: 5405)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0728	PE2	PE5	RUNX1	RUNX1	+1 C to A	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	AGATGGCTTCAGAC A (SEQ ID NO: 5386)	TGTCTGAAGCCATCTCTT CCTCCTGAAAAT (SEQ ID NO: 5387)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0729	PE2	PE5	RUNX1	RUNX1	+1 C to G	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	GGATGGCTTCAGAC A (SEQ ID NO: 5388)	TGTCTGAAGCCATCCCTT CCTCCTGAAAAT (SEQ ID NO: 5389)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0730	PE2	PE5	RUNX1	RUNX1	+1 C to T	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	TGATGGCTTCAGAC A (SEQ ID NO: 5390)	TGTCTGAAGCCATCACTT CCTCCTGAAAAT (SEQ ID NO: 5391)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0731	PE2	PE5	RUNX1	RUNX1	+2 G to A	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CAATGGCTTCAGAC A (SEQ ID NO: 5392)	TGTCTGAAGCCATTGCTT CCTCCTGAAAAT (SEQ ID NO: 5393)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G, 45E,45F,50A, 50B,50G,51B, 51D	HEK293 T	Chen2021 _0732	PE2	PE5	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0733	PE2	PE5	RUNX1	RUNX1	+3 A to G	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGGTGGCTTCAGAC A (SEQ ID NO: 5394)	TGTCTGAAGCCACCGCT TCCTCCTGAAAAT (SEQ ID NO: 5395)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0734	PE2	PE5	RUNX1	RUNX1	+3 A to T	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGTTGGCTTCAGAC A (SEQ ID NO: 5396)	TGTCTGAAGCCAACGCTT CCTCCTGAAAAT (SEQ ID NO: 5397)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G, 45E,45F,50A, 50B,50G,51B, 51D	HEK293 T	Chen2021 _0735	PE2	PE5	RUNX1	RUNX1	+4 T to A	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGAAGGCTTCAGAC A (SEQ ID NO: 5398)	TGTCTGAAGCCTTCGCTT CCTCCTGAAAAT (SEQ ID NO: 5399)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0736	PE2	PE5	RUNX1	RUNX1	+4 T to C	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGACGGCTTCAGAC A (SEQ ID NO: 5400)	TGTCTGAAGCCGTCGCT TCCTCCTGAAAAT (SEQ ID NO: 5401)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293 T	Chen2021 _0737	PE2	PE5	RUNX1	RUNX1	+4 T to G	+38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGAGGGCTTCAGAC A (SEQ ID NO: 5402)	TGTCTGAAGCCCTCGCTT CCTCCTGAAAAT (SEQ ID NO: 5403)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

44A,44B,44G,45E,45F,50A,50B,50C,50G,51B,51D	HEK293T	Chen2021_0738	PE2	PE5	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50C,50G,51B,51D	HEK293T	Chen2021_0739	PE2	PE5	RUNX1	RUNX1	+6 G to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATGCCTTCAGACA (SEQ ID NO: 5404)	TGTCTGAAGGCATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5405)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0740	PE2	PE2	DNMT1	DNMT1	+1 A to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0741	PE2	PE2	DNMT1	DNMT1	+1 A to G		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	GCAGGGGTGAC (SEQ ID NO: 5408)	GTCACCCCTGCTTCTGGCACCAGG (SEQ ID NO: 5409)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0742	PE2	PE2	DNMT1	DNMT1	+2 C to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	AAAGGGGTGAC (SEQ ID NO: 5410)	GTCACCCCTTTTTCTGGCACCAGG (SEQ ID NO: 5411)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0743	PE2	PE2	DNMT1	DNMT1	+2 C to G		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	AGAGGGGTGAC (SEQ ID NO: 5412)	GTCACCCCTCTTTCTGGCACCAGG (SEQ ID NO: 5413)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0744	PE2	PE2	DNMT1	DNMT1	+2 C to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ATAGGGGTGAC (SEQ ID NO: 5414)	GTCACCCCTATTTCTGGCACCAGG (SEQ ID NO: 5415)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0745	PE2	PE2	DNMT1	DNMT1	+3 A to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACTGGGGTGAC (SEQ ID NO: 5416)	GTCACCCCAGTTTCTGGCACCAGG (SEQ ID NO: 5417)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0746	PE2	PE2	DNMT1	DNMT1	+4 G to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAAGGGTGAC (SEQ ID NO: 5418)	GTCACCCTTGTTTCTGGCACCAGG (SEQ ID NO: 5419)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0747	PE2	PE2	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGG (SEQ ID NO: 5421)	0	0	0	0
44A,44B,44G,50A,50B,50C,50G,51D	HEK293T	Chen2021_0748	PE2	PE2	DNMT1	DNMT1	+6 G to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0749	PE2	PE2	DNMT1	DNMT1	+8 T to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGAGACGGGAGGGC (SEQ ID NO: 5424)	GCCCTCCCGTCTCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5425)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0750	PE2	PE2	DNMT1	DNMT1	+8 T to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGCGACGGGAGGGC (SEQ ID NO: 5263)	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5264)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0751	PE2	PE2	DNMT1	DNMT1	+8 T to G		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGGGACGGGAGGGC (SEQ ID NO: 5426)	GCCCTCCCGTCCCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5427)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0752	PE2	PE4	DNMT1	DNMT1	+1 A to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0753	PE2	PE4	DNMT1	DNMT1	+1 A to G		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	GCAGGGGTGAC (SEQ ID NO: 5408)	GTCACCCCTGCTTCTGGCACCAGG (SEQ ID NO: 5409)	0	0	0	0

44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0754	PE2	PE4	DNMT1	DNMT1	+2 C to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	AAAGGGGTGAC (SEQ ID NO: 5410)	GTCACCCCTTTTTCTGGCACCAGG (SEQ ID NO: 5411)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0755	PE2	PE4	DNMT1	DNMT1	+2 C to G		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	AGAGGGGTGAC (SEQ ID NO: 5412)	GTCACCCCTCTTTCTGGCACCAGG (SEQ ID NO: 5413)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0756	PE2	PE4	DNMT1	DNMT1	+2 C to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ATAGGGGTGAC (SEQ ID NO: 5414)	GTCACCCCTATTTCTGGCACCAGG (SEQ ID NO: 5415)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0757	PE2	PE4	DNMT1	DNMT1	+3 A to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACTGGGGTGAC (SEQ ID NO: 5416)	GTCACCCCAGTTTCTGGCACCAGG (SEQ ID NO: 5417)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0758	PE2	PE4	DNMT1	DNMT1	+4 G to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAAGGGTGAC (SEQ ID NO: 5418)	GTCACCCTTGTTTCTGGCACCAGG (SEQ ID NO: 5419)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0759	PE2	PE4	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGG (SEQ ID NO: 5421)	0	0	0	0
44A,44B,44G,50A,50B,50C,50G,51D	HEK293T	Chen2021_0760	PE2	PE4	DNMT1	DNMT1	+6 G to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0761	PE2	PE4	DNMT1	DNMT1	+8 T to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGAGACGGGAGGGC (SEQ ID NO: 5424)	GCCCTCCCGTCTCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5425)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0762	PE2	PE4	DNMT1	DNMT1	+8 T to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGCGACGGGAGGGC (SEQ ID NO: 5263)	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5264)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0763	PE2	PE4	DNMT1	DNMT1	+8 T to G		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGGGACGGGAGGGC (SEQ ID NO: 5426)	GCCCTCCCGTCCCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5427)	0	0	0	0
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0764	PE2	PE3	DNMT1	DNMT1	+1 A to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0765	PE2	PE3	DNMT1	DNMT1	+1 A to G	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	GCAGGGGTGAC (SEQ ID NO: 5408)	GTCACCCCTGCTTCTGGCACCAGG (SEQ ID NO: 5409)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0766	PE2	PE3	DNMT1	DNMT1	+2 C to A	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	AAAGGGGTGAC (SEQ ID NO: 5410)	GTCACCCCTTTTTCTGGCACCAGG (SEQ ID NO: 5411)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0767	PE2	PE3	DNMT1	DNMT1	+2 C to G	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	AGAGGGGTGAC (SEQ ID NO: 5412)	GTCACCCCTCTTTCTGGCACCAGG (SEQ ID NO: 5413)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0768	PE2	PE3	DNMT1	DNMT1	+2 C to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ATAGGGGTGAC (SEQ ID NO: 5414)	GTCACCCCTATTTCTGGCACCAGG (SEQ ID NO: 5415)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0769	PE2	PE3	DNMT1	DNMT1	+3 A to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACTGGGGTGAC (SEQ ID NO: 5416)	GTCACCCCAGTTTCTGGCACCAGG (SEQ ID NO: 5417)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0770	PE2	PE3	DNMT1	DNMT1	+4 G to A	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAAGGGTGAC (SEQ ID NO: 5418)	GTCACCCTTGTTTCTGGCACCAGG (SEQ ID NO: 5419)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50C,50G,51D	HEK293T	Chen2021_0771	PE2	PE3	DNMT1	DNMT1	+5 G to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGG (SEQ ID NO: 5421)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45D,45E,45F,50A,50B,50C,50G,51D	HEK293T	Chen2021_0772	PE2	PE3	DNMT1	DNMT1	+6 G to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0773	PE2	PE3	DNMT1	DNMT1	+8 T to A	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGAGACGGGAGGGC (SEQ ID NO: 5424)	GCCCTCCCGTCTCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5425)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0774	PE2	PE3	DNMT1	DNMT1	+8 T to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGCGACGGGAGGGC (SEQ ID NO: 5263)	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5264)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0775	PE2	PE3	DNMT1	DNMT1	+8 T to G	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGGGACGGGAGGGC (SEQ ID NO: 5426)	GCCCTCCCGTCCCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5427)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0776	PE2	PE5	DNMT1	DNMT1	+1 A to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0777	PE2	PE5	DNMT1	DNMT1	+1 A to G	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	GCAGGGGTGAC (SEQ ID NO: 5408)	GTCACCCCTGCTTCTGGCACCAGG (SEQ ID NO: 5409)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0778	PE2	PE5	DNMT1	DNMT1	+2 C to A	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	AAAGGGGTGAC (SEQ ID NO: 5410)	GTCACCCCTTTTTCTGGCACCAGG (SEQ ID NO: 5411)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0779	PE2	PE5	DNMT1	DNMT1	+2 C to G	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	AGAGGGGTGAC (SEQ ID NO: 5412)	GTCACCCCTCTTTCTGGCACCAGG (SEQ ID NO: 5413)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0780	PE2	PE5	DNMT1	DNMT1	+2 C to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ATAGGGGTGAC (SEQ ID NO: 5414)	GTCACCCCTATTTCTGGCACCAGG (SEQ ID NO: 5415)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0781	PE2	PE5	DNMT1	DNMT1	+3 A to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACTGGGGTGAC (SEQ ID NO: 5416)	GTCACCCCAGTTTCTGGCACCAGG (SEQ ID NO: 5417)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0782	PE2	PE5	DNMT1	DNMT1	+4 G to A	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAAGGGTGAC (SEQ ID NO: 5418)	GTCACCCTTGTTTCTGGCACCAGG (SEQ ID NO: 5419)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50C,50G,51D	HEK293T	Chen2021_0783	PE2	PE5	DNMT1	DNMT1	+5 G to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGG (SEQ ID NO: 5421)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45D,45E,45F,50A,50B,50C,50G,51D	HEK293T	Chen2021_0784	PE2	PE5	DNMT1	DNMT1	+6 G to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0785	PE2	PE5	DNMT1	DNMT1	+8 T to A	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGAGACGGGAGGGC (SEQ ID NO: 5424)	GCCCTCCCGTCTCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5425)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0786	PE2	PE5	DNMT1	DNMT1	+8 T to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGCGACGGGAGGGC (SEQ ID NO: 5263)	GCCCTCCCGTCGCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5264)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45D,45E,45F,50A,50B,50G,51D	HEK293T	Chen2021_0787	PE2	PE5	DNMT1	DNMT1	+8 T to G	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGGGACGGGAGGGC (SEQ ID NO: 5426)	GCCCTCCCGTCCCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5427)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0788	PE2	PE2	VEGFA	VEGFA	+1 T to A		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	AGAGGGCTCCAGATGGCACATT (SEQ ID NO: 5428)	AATGTGCCATCTGGAGCCCTCTTCTGGCCTGCAGA (SEQ ID NO: 5429)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0789	PE2	PE2	VEGFA	VEGFA	+1 T to C		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	CGAGGGCTCCAGATGGCACATT (SEQ ID NO: 5430)	AATGTGCCATCTGGAGCCCTCGTCTGGCCTGCAGA (SEQ ID NO: 5431)	0	0	0	0
44A,44B,44G,50A,50B,50G,51D	HEK293T	Chen2021_0790	PE2	PE2	VEGFA	VEGFA	+1 T to G		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	GGAGGGCTCCAGATGGCACATT (SEQ ID NO: 5432)	AATGTGCCATCTGGAGCCCTCCTCTGGCCTGCAGA (SEQ ID NO: 5433)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0791	PE2	PE2	VEGFA	VEGFA	+2 G to A		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TAAGGGCTCCAGATGGCACATT (SEQ ID NO: 5434)	AATGTGCCATCTGGAGCCCTTATCTGGCCTGCAGA (SEQ ID NO: 5435)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0792	PE2	PE2	VEGFA	VEGFA	+3 A to C		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGCGGGCTCCAGATGGCACATT (SEQ ID NO: 5436)	AATGTGCCATCTGGAGCCCGCATCTGGCCTGCAGA (SEQ ID NO: 5437)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0793	PE2	PE2	VEGFA	VEGFA	+3 A to G		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGGGGGCTCCAGATGGCACATT (SEQ ID NO: 5438)	AATGTGCCATCTGGAGCCCCCATCTGGCCTGCAGA (SEQ ID NO: 5439)	0	0	0	0
44A,44B,50A,50B,50G,51D	HEK293T	Chen2021_0794	PE2	PE2	VEGFA	VEGFA	+3 A to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGTGGGCTCCAGATGGCACATT (SEQ ID NO: 5267)	AATGTGCCATCTGGAGCCCACATCTGGCCTGCAGA (SEQ ID NO: 5268)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0795	PE2	PE2	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAGA (SEQ ID NO: 5441)	0	0	0	0
44A,44B,50A,50B,50C,50G,51D	HEK293T	Chen2021_0796	PE2	PE2	VEGFA	VEGFA	+6 G to C		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGCCTCCAGATGGCACATT (SEQ ID NO: 5442)	AATGTGCCATCTGGAGGCCTCATCTGGCCTGCAGA (SEQ ID NO: 5443)	0	0	0	0

44A,44B,44G, 50A,50B,50G, 51D	HEK293T	Chen2021_0797	PE2	PE2	VEGFA	VEGFA	+7 C to A		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCC CTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	0	0
44A,44B,44G, 50A,50B,50G, 51D	HEK293T	Chen2021_0798	PE2	PE2	VEGFA	VEGFA	+7 C to T		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGTTCCAGAT GGCACATT (SEQ ID NO: 5446)	AATGTGCCATCTGGAAC CCTCATCTGGCCTGCAG A (SEQ ID NO: 5447)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0799	PE2	PE2	VEGFA	VEGFA	+9 C to G		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCTGCAGAT GGCACATT (SEQ ID NO: 5448)	AATGTGCCATCTGCAGC CCTCATCTGGCCTGCAG A (SEQ ID NO: 5449)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0800	PE2	PE4	VEGFA	VEGFA	+1 T to A		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	AGAGGGCTCCAGAT GGCACATT (SEQ ID NO: 5428)	AATGTGCCATCTGGAGC CCTCTTCTGGCCTGCAG A (SEQ ID NO: 5429)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0801	PE2	PE4	VEGFA	VEGFA	+1 T to C		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	CGAGGGCTCCAGAT GGCACATT (SEQ ID NO: 5430)	AATGTGCCATCTGGAGC CCTCGTCTGGCCTGCAG A (SEQ ID NO: 5431)	0	0	0	0
44A,44B,44G, 50A,50B,50G, 51D	HEK293T	Chen2021_0802	PE2	PE4	VEGFA	VEGFA	+1 T to G		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	GGAGGGCTCCAGA TGGCACATT (SEQ ID NO: 5432)	AATGTGCCATCTGGAGC CCTCCTCTGGCCTGCAG A (SEQ ID NO: 5433)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0803	PE2	PE4	VEGFA	VEGFA	+2 G to A		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TAAGGGCTCCAGAT GGCACATT (SEQ ID NO: 5434)	AATGTGCCATCTGGAGC CCTTATCTGGCCTGCAG A (SEQ ID NO: 5435)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0804	PE2	PE4	VEGFA	VEGFA	+3 A to C		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGCGGGCTCCAGAT GGCACATT (SEQ ID NO: 5436)	AATGTGCCATCTGGAGC CCGCATCTGGCCTGCAG A (SEQ ID NO: 5437)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0805	PE2	PE4	VEGFA	VEGFA	+3 A to G		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGGGGGCTCCAGAT GGCACATT (SEQ ID NO: 5438)	AATGTGCCATCTGGAGC CCCCATCTGGCCTGCAG A (SEQ ID NO: 5439)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0806	PE2	PE4	VEGFA	VEGFA	+3 A to T		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGTGGGCTCCAGAT GGCACATT (SEQ ID NO: 5267)	AATGTGCCATCTGGAGC CCACATCTGGCCTGCAG A (SEQ ID NO: 5268)	0	0	0	0
44A,44B,50A, 50B,50C,50G, 51D	HEK293T	Chen2021_0807	PE2	PE4	VEGFA	VEGFA	+5 G to T		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGTGCTCCAGAT GGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGC ACTCATCTGGCCTGCAG A (SEQ ID NO: 5441)	0	0	0	0
44A,44B,50A, 50B,50C,50G, 51D	HEK293T	Chen2021_0808	PE2	PE4	VEGFA	VEGFA	+6 G to C		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGCCTCCAGAT GGCACATT (SEQ ID NO: 5442)	AATGTGCCATCTGGAGG CCTCATCTGGCCTGCAG A (SEQ ID NO: 5443)	0	0	0	0
44A,44B,44G, 50A,50B,50G, 51D	HEK293T	Chen2021_0809	PE2	PE4	VEGFA	VEGFA	+7 C to A		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCC CTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	0	0
44A,44B,44G, 50A,50B,50G, 51D	HEK293T	Chen2021_0810	PE2	PE4	VEGFA	VEGFA	+7 C to T		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGTTCCAGAT GGCACATT (SEQ ID NO: 5446)	AATGTGCCATCTGGAAC CCTCATCTGGCCTGCAG A (SEQ ID NO: 5447)	0	0	0	0
44A,44B,50A, 50B,50G,51D	HEK293T	Chen2021_0811	PE2	PE4	VEGFA	VEGFA	+9 C to G		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCTGCAGAT GGCACATT (SEQ ID NO: 5448)	AATGTGCCATCTGCAGC CCTCATCTGGCCTGCAG A (SEQ ID NO: 5449)	0	0	0	0
44A,44B,45E, 45F,50A,50B, 50G,51B,51D	HEK293T	Chen2021_0812	PE2	PE3	VEGFA	VEGFA	+1 T to A	+57	GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	AGAGGGCTCCAGAT GGCACATT (SEQ ID NO: 5428)	AATGTGCCATCTGGAGC CCTCTTCTGGCCTGCAG A (SEQ ID NO: 5429)	0	0	GATGTACAG AGAGCCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0613	PE2	PE3	VEGFA	VEGFA	+1 T to C	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	CGAGGGCCTCCAGATGGGCACATT (SEQ ID NO: 5430)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5431)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0614	PE2	PE3	VEGFA	VEGFA	+1 T to G	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	GGAGGGCCTCCAGATCCGCACATT (SEQ ID NO: 5432)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5433)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0615	PE2	PE3	VEGFA	VEGFA	+2 G to A	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TAAGGGCCTCCAGATGGCACATT (SEQ ID NO: 5434)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5435)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0616	PE2	PE3	VEGFA	VEGFA	+3 A to C	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TCCGGGCCTCCAGATGCGCACATT (SEQ ID NO: 5436)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5437)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0617	PE2	PE3	VEGFA	VEGFA	+3 A to G	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TGGGGGCCTCCAGATCCCCACATT (SEQ ID NO: 5438)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5439)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0618	PE2	PE3	VEGFA	VEGFA	+3 A to T	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TGTGGGGCTCCAGATGCACACATT (SEQ ID NO: 5267)	9ATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5268)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0619	PE2	PE3	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TGAGTGCTCCAGATGCACACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5441)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0620	PE2	PE3	VEGFA	VEGFA	+6 G to C	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TGAGGGCCTCCAGATGCACACATT (SEQ ID NO: 5442)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5443)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0621	PE2	PE3	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TGAGGGATCCAGATGCACACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTGCAGA (SEQ ID NO: 5445)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0622	PE2	PE3	VEGFA	VEGFA	+7 C to T	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TGAGGGTTCCAGATGCACACATT (SEQ ID NO: 5445)	AATGTGCCATCTGGAACCCTGCAGA (SEQ ID NO: 5447)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0623	PE2	PE3	VEGFA	VEGFA	+9 C to G	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	TGAGGGCCTGCAGATGGCACATT (SEQ ID NO: 5446)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5448)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0624	PE2	PE5	VEGFA	VEGFA	+1 T to A	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	AGAGGGCCTCCAGATGGCACATT (SEQ ID NO: 5426)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5428)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0625	PE2	PE5	VEGFA	VEGFA	+1 T to C	+57	GATGTCTGCAGGCCATGCAGATAAGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCATGCAGAA (SEQ ID NO: 5266)	CGAGGGCCTCCAGATGGCACATT (SEQ ID NO: 5430)	AATGTGCCATCTGGAGCCCTGCAGA (SEQ ID NO: 5431)	0		GATGTACAGGCCATGCAGATAAGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0626	PE2	PE5	VEGFA	VEGFA	+1 T to G	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	CGAGGGCTCCAGATTGGCACATT (SEQ ID NO: 5432)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5433)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0627	PE2	PE5	VEGFA	VEGFA	+2 G to A	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TAAGGGCTCCAGATTGGCACATT (SEQ ID NO: 5434)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5435)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0628	PE2	PE5	VEGFA	VEGFA	+3 A to C	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TGCGGGCTCCAGATGGCACATT (SEQ ID NO: 5436)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5437)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0629	PE2	PE5	VEGFA	VEGFA	+3 A to G	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TGCGGGCTCCAGATGGCACATT (SEQ ID NO: 5438)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5439)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0630	PE2	PE5	VEGFA	VEGFA	+3 A to T	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TGTGGGCTCCAGATGGCACATT (SEQ ID NO: 5267)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5268)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0631	PE2	PE5	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5441)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0632	PE2	PE5	VEGFA	VEGFA	+6 G to C	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TGAGGGCCTCCAGATGGCACATT (SEQ ID NO: 5442)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5443)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0633	PE2	PE5	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TGAGGGAATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5445)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,44G,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0634	PE2	PE5	VEGFA	VEGFA	+7 C to T	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TGAGGGTTCCAGATGGCACATT (SEQ ID NO: 5446)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5447)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44A,44B,45E,45F,50A,50B,50G,51B,51D	HEK293T	Chen2021_0635	PE2	PE5	VEGFA	VEGFA	+9 C to G	+57	GATGTCTGCAGGCCAGACATTAAGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGACATTCCA (SEQ ID NO: 5266)	TGAGGGCTCCAGATGGCACATT (SEQ ID NO: 5448)	AATGTGCCATCTGGAGCCCTCCTCTGCCCAG A (SEQ ID NO: 5449)	0	0	GATGTACAGGACAGAGCCCAGGTGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
50D	HEK293T	Chen2021_0636	PE2	NT siRNA	RNF2	RNF2	+3 G to C		GTCATCTTATGCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTACA (SEQ ID NO: 5252)	AAGCAACACCTGAGGTA (SEQ ID NO: 5372)	0	0	0	0
50D	HEK293T	Chen2021_0637	PE2	MSH2 siRNA	RNF2	RNF2	+3 G to C		GTCATCTTATGCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTACA (SEQ ID NO: 5252)	CTCAAGGTGTTCGTTATGCATTA (SEQ ID NO: 5372)	AAGCAACACCTGAGGTA (SEQ ID NO: 5373)	0	0	0	0
50D	HEK293T	Chen2021_0638	PE2	MSH6 siRNA	RNF2	RNF2	+3 G to C		GTCATCTTATGCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTACA (SEQ ID NO: 5252)	CTCAAGGTGTTCGTTATGCATTA (SEQ ID NO: 5372)	AAGCAACACCTGAGGTA (SEQ ID NO: 5373)	0	0	0	0
50D	HEK293T	Chen2021_0639	PE2	MLH1 siRNA	RNF2	RNF2	+3 G to C		GTCATCTTATGCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTACA (SEQ ID NO: 5252)	CTCAAGGTGTTCGTTATGCATTA (SEQ ID NO: 5372)	AAGCAACACCTGAGGTA (SEQ ID NO: 5373)	0	0	0	0

50D	HEK293T	Chen2021_0840	PE2	PMS2 siRNA	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	0
50D	HEK293T	Chen2021_0841	PE2	NT siRNA	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
50D	HEK293T	Chen2021_0842	PE2	MSH2 siRNA	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
50D	HEK293T	Chen2021_0843	PE2	MSH6 siRNA	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
50D	HEK293T	Chen2021_0844	PE2	MLH1 siRNA	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
50D	HEK293T	Chen2021_0845	PE2	PMS2 siRNA	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
50E	HAP1 wild-type	Chen2021_0846	PE2-P2A-BSD	PE2	RNP2	RNP2	+3 G to C		GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	0	0
50E	HAP1 ΔMSH2	Chen2021_0847	PE2-P2A-BSD	PE2	RNP2	RNP2	+3 G to C		GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	0	0
50E	HAP1 ΔMLH1	Chen2021_0849	PE2-P2A-BSD	PE2	RNP2	RNP2	+3 G to C		GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	0	0
50E	HAP1 wild-type	Chen2021_0849	PE2-P2A-BSD	PE3	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
50E	HAP1 ΔMSH2	Chen2021_0850	PE2-P2A-BSD	PE3	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
50E	HAP1 ΔMLH1	Chen2021_0851	PE2-P2A-BSD	PE3	RNP2	RNP2	+3 G to C	+41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTCAGGTGTTCGTT ATGACTAAGATG (SEQ ID NO: 5373)	AACGAACACCTGAGGTTA (SEQ ID NO: 5373)	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
50F	HeLa CRISPRi cells with integrated pPC1000 inch-	Chen2021_0852	SpPE 2-P2A-BSD	PE2_NT CRISPRi sgRNA	pPC1000 edit site	pPC1000 9 edit site	+6 G to A		GTCTGATAGGC AGCCTGCACTG (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTAGAG TAATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGAGGTGAATT C (SEQ ID NO: 5450)	GAATTCAGCTCACCGAGTGG CAGGCT (SEQ ID NO: 5451)	0	0	0	0

50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0653	SaPE 2-P2 A-BS D	PE2, NT CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to C		GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTCACCGCACCAGT GCAGGCT (SEQ ID NO: 5186)	0	0	0	0
50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0654	SaPE 2-P2 A-BS D	PE2, NT CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to T		GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGTGGTGAATT C (SEQ ID NO: 5452)	GAATTCACCACACCAGTGT CAGGCT (SEQ ID NO: 5453)	0	0	0	0
50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0655	SaPE 2-P2 A-BS D	PE3+50, NT CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to A	+50	GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGAGGTGAATT C (SEQ ID NO: 5450)	GAATTCACCTCACCAGTGT CAGGCT (SEQ ID NO: 5451)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTGTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5188)
50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0656	SaPE 2-P2 A-BS D	PE3+50, NT CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to C	+50	GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGCGTGCGGTGAATT C (SEQ ID NO: 5185)	GAATTCACCGCACCAGTGT GCAGGCT (SEQ ID NO: 5186)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTGTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5188)
50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0657	SaPE 2-P2 A-BS D	PE3+50, NT CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to T	+50	GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGTGGTGAATT C (SEQ ID NO: 5452)	GAATTCACCACACCAGTGT CAGGCT (SEQ ID NO: 5453)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTGTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5188)
50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0658	SaPE 2-P2 A-BS D	PE2, MSH2-targetting CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to A		GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGAGGTGAATT C (SEQ ID NO: 5450)	GAATTCACCTCACCAGTGT CAGGCT (SEQ ID NO: 5451)	0	0	0	0
50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0659	SaPE 2-P2 A-BS D	PE2, MSH2-targetting CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to C		GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGCGGTGAATT C (SEQ ID NO: 5195)	GAATTCACCGCACCAGT GCAGGCT (SEQ ID NO: 5196)	0	0	0	0
50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0660	SaPE 2-P2 A-BS D	PE2, MSH2-targetting CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to T		GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGTGGTGAATT C (SEQ ID NO: 5452)	GAATTCACCACACCAGTGT CAGGCT (SEQ ID NO: 5453)	0	0	0	0
50F	HeLa CRISPR cells with integrated pPC1000 inch-	Chen2021 0661	SaPE 2-P2 A-BS D	PE3+50, MSH2-targetting CRISPR sgRNA	pPC1000 edit site	pPC100 0 edit site	+6 G to A	+50	GTCTGATAGGC AGCCTGCACTG G (SEQ ID NO: 5182)	GTTTTAGTACTCTGTAATGAAAATTACAG AATCTACTAAAACAAGGCAAAATGCCGT GTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCAC (SEQ ID NO: 5184)	TGGTGAGGTGAATT C (SEQ ID NO: 5450)	GAATTCACCTCACCAGTGT CAGGCT (SEQ ID NO: 5451)	0	0	GGTTCAACA AAATATTGT CGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAA TTACAGAATCTACTAAAACAAGG CAAAATGCCGTGTTTATCTCGTC AACTTGTTGGCGAGA (SEQ ID NO: 5188)

50F	HeLa CRISPRi cells with integrated pPC1000 insti-	Chen2021_0662	SaPE2-P2 A–BS ID	PE3+50 MSH2-targeting CRISPR sgRNA	pPC1000 edit site	pPC1000 6 edit site	+6 G to T	+50	GTCTGATAGGCAGCCTGCACTTTATTAATGAAAATTACAGAGTTTATAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5182)	GTCTGATAGGCGTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)	AGCCTGCACTGCCTGCACTTTATTAATGAAAATTACAGAGTTTATAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5184)	TCGTCGTGCACTGCCTGCACTTTATTAATGAAAATTACAGAGTTTATAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5185)	GAATTGACGGTGCAGCAGGTCGAATTTGAATTGACGGTGCAGCAGGTCGCAGGCGTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5187)	GTTTCAACAAAAATATTGTCGAG (SEQ ID NO: 5187)	GTTTTAGTACTCTGTAATGAAAATTACAGAATCTACTAAAACAAGGCAAAATGCCGTGTTTTATCTCGTCAACTTGTTGGCGAGA (SEQ ID NO: 5183)
44C.50H	HEK293T	Chen2021_0654	PE2	HEK3	HEK3	+1 A ins		GGCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5231)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	ATGATGGCGAGA (SEQ ID NO: 5454)	TCTGCCATCATCGTGCTCAAGTCTG (SEQ ID NO: 5455)	0	0	0
44C.50H	HEK293T	Chen2021_0655	PE2	HEK3	HEK3	+1 CTT ins		GGCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5231)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	CTTTGATGGCGAGA (SEQ ID NO: 5456)	TCTGCCATCAAAGCGTGCTCTGCCATCATCGTGCTCAAGTCTG (SEQ ID NO: 5457)	0	0	0
44C.50H	HEK293T	Chen2021_0656	PE2	HEK3	HEK3	+1 T del		GGCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5231)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	GATGGCGAGA (SEQ ID NO: 5458)	TCTGCCATCGGTGCTCAAGTCTG (SEQ ID NO: 5459)	0	0	0
44C.50H	HEK293T	Chen2021_0657	PE2	HEK3	HEK3	+1–3 TGA del		GGCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	TGGCAGAGGAAGC (SEQ ID NO: 5460)	TGGAAGGAAGCAGGGCTTCGCCTTCCTCTGCTCCAGCGTGCTCAAGTCTG (SEQ ID NO: 5461)	0	0	0
44C.50H	HEK293T	Chen2021_0658	PE2	HEK3	HEK3	+1 A ins		GGCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5231)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	ATGATGGCGAGA (SEQ ID NO: 5454)	TCTGCCATCATCGTGCTCAAGTCTG (SEQ ID NO: 5455)	0	0	0
44C.50H	HEK293T	Chen2021_0659	PE2	HEK3	HEK3	+1 CTT ins		GGCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5231)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	CTTTGATGGCGAGA (SEQ ID NO: 5456)	TCTGCCATCAAAGCGTGCTCTGCCATCATCGTGCTCAAGTCTG (SEQ ID NO: 5457)	0	0	0
44C.50H	HEK293T	Chen2021_0670	PE2	HEK3	HEK3	+1 T del		GCCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5239)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	GATGGCAGCAGG (SEQ ID NO: 5458)	TCTGCCATCGGTGCTCAAGTCTG (SEQ ID NO: 5459)	0	0	0
44C.50H	HEK293T	Chen2021_0671	PE2	HEK3	HEK3	+1–3 TGA del		GCCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	TCCGCAGAGGAAGCAGGG (SEQ ID NO: 5460)	TGGAGGAAGGCAGGGCTTCGCCTTCCTCTGCTCCAGCGTGCTCAAGTCTG (SEQ ID NO: 5461)	0	0	0
44C	HEK293T	Chen2021_0672	PE3	HEK3	HEK3	+1 A ins	+90	GCCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	ATGATGGCGAGA (SEQ ID NO: 5454)	TCTGCCATCATCGTGCTCAAGTCTG (SEQ ID NO: 5455)	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0673	PE3	HEK3	HEK3	+1 CTT ins	+90	GCCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	CTTTGATGGCGAGA (SEQ ID NO: 5456)	TCTGCCATCAAAGCGTGCTCTGCCATCATCGTGCTCAAGTCTG (SEQ ID NO: 5457)	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0674	PE3	HEK3	HEK3	+1 T del	+90	GGCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	GATGGCGAGA (SEQ ID NO: 5460)	TCTGCCATCGGTGCTCAAGTCTG (SEQ ID NO: 5459)	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0675	PE3	HEK3	HEK3	+1–3 TGA del	+90	GGCCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGCGCCAGACTGAGCACGTGAAGCGTGGC (SEQ ID NO: 5232)	TGGAGGAAGGCAGGGCTTCCTCTGCTCCAGCGTGCTCAAGTCTG (SEQ ID NO: 5461)	TGGAGGAAGGCAGGGCTTCGCCTTCCTCTGCTCCAGCGTGCTCAAGTCTG (SEQ ID NO: 5461)	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCTTTTTT (SEQ ID NO: 5237)

44C	HEK293T	Chen2021_0876	PE2	PE5	HEK3	HEK3	+1 A ins	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATGATGGCAGA (SEQ ID NO: 5454)	TCTGCCATCATCGTGCTCAGTCTG (SEQ ID NO: 5455)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0877	PE2	PE5	HEK3	HEK3	+1 CTT ins	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	CTTTGATGGCAGA (SEQ ID NO: 5456)	TCTGCCATCAAAGCGTGCTCAGTCTG (SEQ ID NO: 5457)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0878	PE2	PE5	HEK3	HEK3	+1 T del	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATGGCAGA (SEQ ID NO: 5458)	TCTGCCATCCGTGCTCAGTCTG (SEQ ID NO: 5459)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0879	PE2	PE5	HEK3	HEK3	+1–3 TGA del	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5460)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCACGTGCTCAGTCTG (SEQ ID NO: 5461)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C,50H	HEK293T	Chen2021_0880	PE2	PE2	EMX1	EMX1	+6 T ins		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGTGCTCCCATCAC (SEQ ID NO: 5462)	GTGATGGGAGCACCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5463)	0	0	0	0
44C,50H	HEK293T	Chen2021_0881	PE2	PE2	EMX1	EMX1	+1 TGC ins		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	TGCGAAGGGCTCCCATCAC (SEQ ID NO: 5464)	GTGATGGGAGCCCTTCGCATTCTTCTGCTCGGA (SEQ ID NO: 5465)	0	0	0	0
44C,50H	HEK293T	Chen2021_0882	PE2	PE2	EMX1	EMX1	+5 G del		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGCTCCCATCAC (SEQ ID NO: 5466)	GTGATGGGAGCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5467)	0	0	0	0
44C,50H	HEK293T	Chen2021_0883	PE2	PE2	EMX1	EMX1	+4–6 GGG del		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAACTCCCATCAC (SEQ ID NO: 5468)	GTGATGGGAGTTCTTCTTCTGCTCGGA (SEQ ID NO: 5469)	0	0	0	0
44C,50H	HEK293T	Chen2021_0884	PE2	PE4	EMX1	EMX1	+6 T ins		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGTGCTCCCATCAC (SEQ ID NO: 5462)	GTGATGGGAGCACCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5463)	0	0	0	0
44C,50H	HEK293T	Chen2021_0885	PE2	PE4	EMX1	EMX1	+1 TGC ins		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	TGCGAAGGGCTCCCATCAC (SEQ ID NO: 5464)	GTGATGGGAGCCCTTCGCATTCTTCTGCTCGGA (SEQ ID NO: 5465)	0	0	0	0
44C,50H	HEK293T	Chen2021_0886	PE2	PE4	EMX1	EMX1	+5 G del		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGCTCCCATCAC (SEQ ID NO: 5466)	GTGATGGGAGCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5467)	0	0	0	0
44C,50H	HEK293T	Chen2021_0887	PE2	PE4	EMX1	EMX1	+4–6 GGG del		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAACTCCCATCAC (SEQ ID NO: 5468)	GTGATGGGAGTTCTTCTTCTGCTCGGA (SEQ ID NO: 5469)	0	0	0	0
44C	HEK293T	Chen2021_0888	PE2	PE3	EMX1	EMX1	+6 T ins	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGTGCTCCCATCAC (SEQ ID NO: 5462)	GTGATGGGAGCACCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5463)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0889	PE2	PE3	EMX1	EMX1	+1 TGC ins	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	TGCGAAGGGCTCCCATCAC (SEQ ID NO: 5464)	GTGATGGGAGCCCTTCGCATTCTTCTGCTCGGA (SEQ ID NO: 5465)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0890	PE2	PE3	EMX1	EMX1	+5 G del	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGCTCCCATCAC (SEQ ID NO: 5466)	GTGATGGGAGCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5467)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44C	HEK293T	Chen2021_0891	PE2	PE3	EMX1	EMX1	+4-6 GGG del	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAACTCCCATCAC (SEQ ID NO: 5468)	GTGATGGGAGTTCTTCTTCTGCTCGGA (SEQ ID NO: 5469)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0892	PE2	PE5	EMX1	EMX1	+6 T ins	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGTGCTCCCATCAC (SEQ ID NO: 5462)	GTGATGGGAGCACCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5463)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0893	PE2	PE5	EMX1	EMX1	+1 TGC ins	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	TGCGAAGGGCTCCCATCAC (SEQ ID NO: 5464)	GTGATGGGAGCCCTTCGCATTCTTCTGCTCGGA (SEQ ID NO: 5465)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0894	PE2	PE5	EMX1	EMX1	+5 G del	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGCTCCCATCAC (SEQ ID NO: 5466)	GTGATGGGAGCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5467)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0895	PE2	PE5	EMX1	EMX1	+4-6 GGG del	-57	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAACTCCCATCAC (SEQ ID NO: 5468)	GTGATGGGAGTTCTTCTTCTGCTCGGA (SEQ ID NO: 5469)	0	0	GCCGTTTGTACTTTGTCCTC (SEQ ID NO: 5241)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0896	PE2	PE2	FANCF	FANCF	+3 C ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCCTGGATCGCTTTTCC (SEQ ID NO: 5470)	GGAAAAGCGATCCAGGGTGCTGCAGAAGGGAT (SEQ ID NO: 5471)	0	0	0	0
44C	HEK293T	Chen2021_0897	PE2	PE2	FANCF	FANCF	+4 GAT ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGATTGGATCGCTTTTCC (SEQ ID NO: 5472)	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5473)	0	0	0	0
44C,50H	HEK293T	Chen2021_0898	PE2	PE2	FANCF	FANCF	+6 G del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGATCGCTTTTCC (SEQ ID NO: 5474)	GGAAAAGCGATCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5475)	0	0	0	0
44C,50H	HEK293T	Chen2021_0899	PE2	PE2	FANCF	FANCF	+5-7 GGA del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTTCGCTTTTCC (SEQ ID NO: 5476)	GGAAAAGCGAAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5477)	0	0	0	0
44C	HEK293T	Chen2021_0900	PE2	PE4	FANCF	FANCF	+3 C ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCCTGGATCGCTTTTCC (SEQ ID NO: 5470)	GGAAAAGCGATCCAGGGTGCTGCAGAAGGGAT (SEQ ID NO: 5471)	0	0	0	0
44C	HEK293T	Chen2021_0901	PE2	PE4	FANCF	FANCF	+4 GAT ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGATTGGATCGCTTTTCC (SEQ ID NO: 5472)	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5473)	0	0	0	0
44C,50H	HEK293T	Chen2021_0902	PE2	PE4	FANCF	FANCF	+6 G del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGATCGCTTTTCC (SEQ ID NO: 5474)	GGAAAAGCGATCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5475)	0	0	0	0
44C,50H	HEK293T	Chen2021_0903	PE2	PE4	FANCF	FANCF	+5-7 GGA del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTTCGCTTTTCC (SEQ ID NO: 5476)	GGAAAAGCGAAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5477)	0	0	0	0
44C	HEK293T	Chen2021_0904	PE2	PE3	FANCF	FANCF	+3 C ins	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCCTGGATCGCTTTTCC (SEQ ID NO: 5470)	GGAAAAGCGATCCAGGGTGCTGCAGAAGGGAT (SEQ ID NO: 5471)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0905	PE2	PE3	FANCF	FANCF	+4 GAT ins	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGATTGGATCGCTTTTCC (SEQ ID NO: 5472)	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5473)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44C	HEK293T	Chen2021_0906	PE2	PE3	FANCF	FANCF	+6 G del	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGATCGCTTTTCC (SEQ ID NO: 5474)	GGAAAAGCGATCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5475)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0907	PE2	PE3	FANCF	FANCF	+5–7 GGA del	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTTCGCTTTTCC (SEQ ID NO: 5476)	GGAAAAGCGAAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5477)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0908	PE2	PE5	FANCF	FANCF	+3 C ins	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCCTGGATCGCTTTTCC (SEQ ID NO: 5470)	GGAAAAGCGATCCAGGGTGCTGCAGAAGGGAT (SEQ ID NO: 5471)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0909	PE2	PE5	FANCF	FANCF	+4 GAT ins	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGATTGGATCGCTTTTCC (SEQ ID NO: 5472)	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5473)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0910	PE2	PE5	FANCF	FANCF	+6 G del	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGATCGCTTTTCC (SEQ ID NO: 5474)	GGAAAAGCGATCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5475)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44C	HEK293T	Chen2021_0911	PE2	PE5	FANCF	FANCF	+5–7 GGA del	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTTCGCTTTTCC (SEQ ID NO: 5476)	GGAAAAGCGAAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5477)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
50H	HEK293T	Chen2021_0912	PE2	PE2	FANCF	FANCF	+1 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5478)	GCCAGAAGCTCGGAAAAGCGATCCAGGGCTGCAGAAGGGAT (SEQ ID NO: 5479)	0	0	0	0
50H	HEK293T	Chen2021_0913	PE2	PE2	FANCF	FANCF	+1–3 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	TGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5480)	GCCAGAAGCTCGGAAAAGCGATCCAGCTGCAGAAGGGAT (SEQ ID NO: 5481)	0	0	0	0
50H	HEK293T	Chen2021_0914	PE2	PE2	FANCF	FANCF	+1–6 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ATCGCTTTTCCGAGCTTCTGGCGGTCT (SEQ ID NO: 5482)	AGACCGCCAGAAGCTCGGAAAAGCGATGCTGCAGAAGGGAT (SEQ ID NO: 5483)	0	0	0	0
50H	HEK293T	Chen2021_0915	PE2	PE2	FANCF	FANCF	+1–10 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CTTTTCCGAGCTTCTGGCGGTCTCAA (SEQ ID NO: 5484)	TTGAGACCGCCAGAAGCTCGGAAAAGGCTGCAGAAGGGAT (SEQ ID NO: 5485)	0	0	0	0
50H	HEK293T	Chen2021_0916	PE2	PE2	FANCF	FANCF	+1–15 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CCGAGCTTCTGGCGGTCTCAAGCACTA (SEQ ID NO: 5486)	TAGTGCTTGAGACCGCCAGAAGCTCGGGCTGCAGAAGGGAT (SEQ ID NO: 5487)	0	0	0	0
50H	HEK293T	Chen2021_0917	PE2	PE2	FANCF	FANCF	+1–20 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CTTCTGGCGGTCTCAAGCACTACCTAC (SEQ ID NO: 5488)	GTAGGTAGTGCTTGAGACCGCCAGAAGGCTGCAGAAGGGAT (SEQ ID NO: 5489)	0	0	0	0
50H	HEK293T	Chen2021_0918	PE2	PE4	FANCF	FANCF	+1 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5478)	GCCAGAAGCTCGGAAAAGCGATCCAGGGCTGCAGAAGGGAT (SEQ ID NO: 5479)	0	0	0	0
50H	HEK293T	Chen2021_0919	PE2	PE4	FANCF	FANCF	+1–3 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	TGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5480)	GCCAGAAGCTCGGAAAAGCGATCCAGCTGCAGAAGGGAT (SEQ ID NO: 5481)	0	0	0	0
50H	HEK293T	Chen2021_0920	PE2	PE4	FANCF	FANCF	+1–6 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ATCGCTTTTCCGAGCTTCTGGCGGTCT (SEQ ID NO: 5482)	AGACCGCCAGAAGCTCGGAAAAGCGATGCTGCAGAAGGGAT (SEQ ID NO: 5483)	0	0	0	0

50H	HEK293T	Chen2021_0921	PE2	PE4	FANCF	FANCF	+1–10 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CTTTTCCGAGCTTCTGGCGGTCTCAA (SEQ ID NO: 5484)	TTGAGACCGCCAGAAGCTCGGAAAAGGCTGCAGAAGGGAT (SEQ ID NO: 5485)	0	0	0	0
50H	HEK293T	Chen2021_0922	PE2	PE4	FANCF	FANCF	+1–15 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CCGAGCTTCTGGCGGTCTCAAGCACTA (SEQ ID NO: 5486)	TAGTGCTTGAGACCGCCAGAAGCTCGGGCTGCAGAAGGGAT (SEQ ID NO: 5487)	0	0	0	0
50H	HEK293T	Chen2021_0923	PE2	PE4	FANCF	FANCF	+1–20 del		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CTTCTGGCGGTCTCAAGCACTACCTAC (SEQ ID NO: 5488)	GTAGGTAGTGCTTGAGACCGCCAGAAGGCTGCAGAAGGGAT (SEQ ID NO: 5489)	0	0	0	0
50H	HEK293T	Chen2021_0924	PE2	PE2	FANCF	FANCF	+3 C ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCCTGGATCGCTTTTCC (SEQ ID NO: 5470)	GGAAAAGCGATCCAGGGTGCTGCAGAAGGGAT (SEQ ID NO: 5471)	0	0	0	0
50H	HEK293T	Chen2021_0925	PE2	PE2	FANCF	FANCF	+4 GAT ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGATTGGATCGCTTTTCC (SEQ ID NO: 5472)	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5473)	0	0	0	0
50H	HEK293T	Chen2021_0926	PE2	PE2	FANCF	FANCF	+1 CACCAT ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CACCATACCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5490)	GCCAGAAGCTCGGAAAAGCGATCCAGGTATGGTGGCTGCAGAAGGGAT (SEQ ID NO: 5491)	0	0	0	0
50H	HEK293T	Chen2021_0927	PE2	PE2	FANCF	FANCF	+1 CACCATCATC ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CACCATCATCACCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5492)	GCCAGAAGCTCGGAAAAGCGATCCAGGTGATGATGGTGGCTGCAGAAGGGAT (SEQ ID NO: 5493)	0	0	0	0
50H	HEK293T	Chen2021_0928	PE2	PE2	FANCF	FANCF	+1 CACCATCATCACCAT ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CACCATCATCACCATACCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5494)	GCCAGAAGCTCGGAAAAGCGATCCAGGTATGGTGATGATGGTGGCTGCAGAAGGGAT (SEQ ID NO: 5495)	0	0	0	0
50H	HEK293T	Chen2021_0929	PE2	PE2	FANCF	FANCF	+1 CACCATCATCACCATCATTA ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CACCATCATCACCATCATTAACCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5496)	GCCAGAAGCTCGGAAAAGCGATCCAGGTTAATGATGGTGATGATGGTGGCTGCAGAAGGGAT (SEQ ID NO: 5497)	0	0	0	0
50H	HEK293T	Chen2021_0930	PE2	PE4	FANCF	FANCF	+3 C ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCCTGGATCGCTTTTCC (SEQ ID NO: 5470)	GGAAAAGCGATCCAGGGTGCTGCAGAAGGGAT (SEQ ID NO: 5471)	0	0	0	0
50H	HEK293T	Chen2021_0931	PE2	PE4	FANCF	FANCF	+4 GAT ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGATTGGATCGCTTTTCC (SEQ ID NO: 5472)	GGAAAAGCGATCCAATCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5473)	0	0	0	0
50H	HEK293T	Chen2021_0932	PE2	PE4	FANCF	FANCF	+1 CACCAT ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CACCATACCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5490)	GCCAGAAGCTCGGAAAAGCGATCCAGGTATGGTGGCTGCAGAAGGGAT (SEQ ID NO: 5491)	0	0	0	0
50H	HEK293T	Chen2021_0933	PE2	PE4	FANCF	FANCF	+1 CACCATCATC ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CACCATCATCACCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5492)	GCCAGAAGCTCGGAAAAGCGATCCAGGTGATGATGGTGGCTGCAGAAGGGAT (SEQ ID NO: 5493)	0	0	0	0
50H	HEK293T	Chen2021_0934	PE2	PE4	FANCF	FANCF	+1 CACCATCATCACCAT ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CACCATCATCACCATACCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO. 5494)	GCCAGAAGCTCGGAAAAGCGATCCAGGTATGGTGATGATGGTGGCTGCAGAAGGGAT (SEQ ID NO. 5495)	0	0	0	0
50H	HEK293T	Chen2021_0935	PE2	PE4	FANCF	FANCF	+1 CACCATCATCACCATCATTA ins		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	CACCATCATCACCATCATTAACCTGGATCGCTTTTCCGAGCTTCTGGC (SEQ ID NO: 5496)	GCCAGAAGCTCGGAAAAGCGATCCAGGTTAATGATGGTGATGATGGTGGCTGCAGAAGGGAT (SEQ ID NO: 5497)	0	0	0	0
50H	HEK293T	Chen2021_0936	PE2	PE2	HEK3	HEK3	+1 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATGGCAGAGGAAAGGAAGCCCT (SEQ ID NO: 5498)	AGGGCTTCCTTTCCTCTGCCATCCGTGCTCAGTCTG (SEQ ID NO: 5499)	0	0	0	0

50H	HEK293T	Chen2021_0937	PE2	PE2	HEK3	HEK3	+1–3 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5460)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCACGTGCTCAGTCTG (SEQ ID NO: 5461)	0	0	0	0	
50H	HEK293T	Chen2021_0938	PE2	PE2	HEK3	HEK3	+1–6 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	CAGAGGAAAGGAAGCCCTGCTTC (SEQ ID NO: 5500)	GAAGCAGGGCTTCCTTTCCTCTGCGTGCTCAGTCTG (SEQ ID NO. 5501)	0	0	0	0	
50H	HEK293T	Chen2021_0939	PE2	PE2	HEK3	HEK3	+1–10 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGAAAGGAAGCCCTGCTTCCTCC (SEQ ID NO: 5502)	GGAGGAAGCAGGGCTTCCTTTCCCGTGCTCAGTCTG (SEQ ID NO: 5503)	0	0	0	0	
50H	HEK293T	Chen2021_0940	PE2	PE2	HEK3	HEK3	+1–15 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGAAGCCCTGCTTCCTCCAGA (SEQ ID NO: 5504)	TCTGGAGGAAGCAGGGCTTCCCGTGCTCAGTCTG (SEQ ID NO: 5505)	0	0	0	0	
50H	HEK293T	Chen2021_0941	PE2	PE2	HEK3	HEK3	+1–20 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	CCCTGCTTCCTCCAGAGGGCGTC (SEQ ID NO: 5506)	GACGCCCTCTGGAGGAAGCAGGGCGTGCTCAGTCTG (SEQ ID NO: 5507)	0	0	0	0	
50H	HEK293T	Chen2021_0942	PE2	PE4	HEK3	HEK3	+1 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATGGCAGAGGAAAGGAAGCCCT (SEQ ID NO: 5498)	AGGGCTTCCTTTCCTCTGCCATCCGTGCTCAGTCTG (SEQ ID NO: 5499)	0	0	0	0	
50H	HEK293T	Chen2021_0943	PE2	PE4	HEK3	HEK3	+1–3 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5460)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCACGTGCTCAGTCTG (SEQ ID NO: 5461)	0	0	0	0	
50H	HEK293T	Chen2021_0944	PE2	PE4	HEK3	HEK3	+1–6 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	CAGAGGAAAGGAAGCCCTGCTTC (SEQ ID NO: 5500)	GAAGCAGGGCTTCCTTTCCTCTGCGTGCTCAGTCTG (SEQ ID NO: 5501)	0	0	0	0	
50H	HEK293T	Chen2021_0945	PE2	PE4	HEK3	HEK3	+1–10 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGAAAGGAAGCCCTGCTTCCTCC (SEQ ID NO: 5502)	GGAGGAAGCAGGGCTTCCTTTCCCGTGCTCAGTCTG (SEQ ID NO: 5503)	0	0	0	0	
50H	HEK293T	Chen2021_0946	PE2	PE4	HEK3	HEK3	+1–15 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGAAGCCCTGCTTCCTCCAGA (SEQ ID NO: 5504)	TCTGGAGGAAGCAGGGCTTCCCGTGCTCAGTCTG (SEQ ID NO: 5505)	0	0	0	0	
50H	HEK293T	Chen2021_0947	PE2	PE4	HEK3	HEK3	+1–20 del		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	CCCTGCTTCCTCCAGAGGGCGTC (SEQ ID NO: 5506)	GACGCCCTCTGGAGGAAGCAGGGCGTGCTCAGTCTG (SEQ ID NO: 5507)	0	0	0	0	
50H	HEK293T	Chen2021_0948	PE2	PE2	HEK3	HEK3	+1 G ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5508)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACCGTGCTCAGTCTG (SEQ ID NO: 5509)	0	0	0	0	
50H	HEK293T	Chen2021_0949	PE2	PE2	HEK3	HEK3	+1 GAT ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5510)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAATCCGTGCTCAGTCTG (SEQ ID NO: 5511)	0	0	0	0	
50H	HEK293T	Chen2021_0950	PE2	PE2	HEK3	HEK3	+1 GATTAC ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTACTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO. 5512)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAGTAATCCGTGCTCAGTCTG (SEQ ID NO: 5513)	0	0	0	0	
50H	HEK293T	Chen2021_0951	PE2	PE2	HEK3	HEK3	+1 GATTACAAGG ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTACAAGGTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5514)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACCTTGTAATCCGTGCTCAGTCTG (SEQ ID NO: 5515)	0	0	0	0	
50H	HEK293T	Chen2021_0952	PE2	PE2	HEK3	HEK3	+1 GATTACAAGGATGAC ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTACAAGGATGACTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5516)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAGTCATCCTTGTAATCCGTGCTCAGTCTG (SEQ ID NO: 5517)	0	0	0	0	

50H	HEK293T	Chen2021_0953	PE2	PE2	HEK3	HEK3	+1 GATTACAAGGATGACGACGA ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTACAAGGATGACGACGATGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5518)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCATCGTCGTCATCCTTGTAATCCGTGCTCAGTCTG (SEQ ID NO: 5519)	0	0	0	0
50H	HEK293T	Chen2021_0954	PE2	PE4	HEK3	HEK3	+1 G ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5508)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACCGTGCTCAGTCTG (SEQ ID NO: 5509)	0	0	0	0
50H	HEK293T	Chen2021_0955	PE2	PE4	HEK3	HEK3	+1 GAT ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5510)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAATCCGTGCTCAGTCTG (SEQ ID NO: 5511)	0	0	0	0
50H	HEK293T	Chen2021_0956	PE2	PE4	HEK3	HEK3	+1 GATTAC ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTACTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5512)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAGTAATCCGTGCTCAGTCTG (SEQ ID NO: 5513)	0	0	0	0
50H	HEK293T	Chen2021_0957	PE2	PE4	HEK3	HEK3	+1 GATTACAAGG ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTACAAGGTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5514)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCACCTTGTAATCCGTGCTCAGTCTG (SEQ ID NO: 5515)	0	0	0	0
50H	HEK293T	Chen2021_0958	PE2	PE4	HEK3	HEK3	+1 GATTACAAGGATGAC ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTACAAGGATGACTGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5516)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCAGTCATCCTTGTAATCCGTGCTCAGTCTG (SEQ ID NO: 5517)	0	0	0	0
50H	HEK293T	Chen2021_0959	PE2	PE4	HEK3	HEK3	+1 GATTACAAGGATGACGACGA ins		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	GATTACAAGGATGACGACGATGATGGCAGAGGAAAGGAAGCCCTGCTTCCTCCA (SEQ ID NO: 5518)	TGGAGGAAGCAGGGCTTCCTTTCCTCTGCCATCATCGTCGTCATCCTTGTAATCCGTGCTCAGTCTG (SEQ ID NO: 5519)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0960	PE2	PE2	DNMT1	DNMT1	+7 G to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGATGACGGGAGGGCAGAAC (SEQ ID NO: 5520)	GTTCTGCCCTCCCGTCATCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5521)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0961	PE2	PE2	DNMT1	DNMT1	+7–8 GT to AC		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGACGACGGGAGGGCAGAAC (SEQ ID NO: 5522)	GTTCTGCCCTCCCGTCGTCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5523)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0962	PE2	PE2	DNMT1	DNMT1	+7–9 GTG to ACC		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGACCACGGGAGGGCAGAAC (SEQ ID NO: 5524)	GTTCTGCCCTCCCGTGGTCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5525)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0963	PE2	PE2	DNMT1	DNMT1	+7–10 GTGA to ACCT		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGACCTCGGGAGGGCAGAAC (SEQ ID NO: 5526)	GTTCTGCCCTCCCGAGGTCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5527)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0964	PE2	PE2	DNMT1	DNMT1	+7–11 GTGAC to ACCTT		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGACCTTGGGAGGGCAGAAC (SEQ ID NO: 5528)	GTTCTGCCCTCCCAAGGTCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5529)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0965	PE2	PE4	DNMT1	DNMT1	+7 G to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGATGACGGGAGGGCAGAAC (SEQ ID NO: 5520)	GTTCTGCCCTCCCGTCATCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5521)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0966	PE2	PE4	DNMT1	DNMT1	+7–8 GT to AC		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGACGACGGGAGGGCAGAAC (SEQ ID NO: 5522)	GTTCTGCCCTCCCGTCGTCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5523)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0967	PE2	PE4	DNMT1	DNMT1	+7–9 GTG to ACC		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGACCACGGGAGGGCAGAAC (SEQ ID NO: 5524)	GTTCTGCCCTCCCGTGGTCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5525)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0968	PE2	PE4	DNMT1	DNMT1	+7–10 GTGA to ACCT		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGACCTCGGGAGGGCAGAAC (SEQ ID NO: 5526)	GTTCTGCCCTCCCGAGGTCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5527)	0	0	0	0

44D,44E,50I	HEK293T	Chen2021_0969	PE2	PE4	DNMT1	DNMT1	+7–11 GTGAC to ACCTT		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGACCTTGGGAGGGCAGAAC (SEQ ID NO: 5528)	GTTCTGCCCTCCCAAGGTCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5529)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0970	PE2	PE2	FANCF	FANCF	+3 C to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGTGGATCGCTTTTCCGAGC (SEQ ID NO: 5530)	GCTCGGAAAAGCGATCCACGTGCTGCAGAAGGGAT (SEQ ID NO: 5531)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0971	PE2	PE2	FANCF	FANCF	+3-4 CT to GA		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGAGGATCGCTTTTCCGAGC (SEQ ID NO: 5532)	GCTCGGAAAAGCGATCCTCGTGCTGCAGAAGGGAT (SEQ ID NO: 5533)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0972	PE2	PE2	FANCF	FANCF	+3-5 CTG to GAA		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGAAGATCGCTTTTCCGAGC (SEQ ID NO: 5534)	GCTCGGAAAAGCGATCTTCGTGCTGCAGAAGGGAT (SEQ ID NO: 5535)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0973	PE2	PE2	FANCF	FANCF	+3-6 CTGG to GAAT		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGAATATCGCTTTTCCGAGC (SEQ ID NO: 5536)	GCTCGGAAAAGCGATATTCGTGCTGCAGAAGGGAT (SEQ ID NO: 5537)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0974	PE2	PE2	FANCF	FANCF	+3-7 CTGGA to GAATC		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGAATCTCGCTTTTCCGAGC (SEQ ID NO: 5538)	GCTCGGAAAAGCGAGATTCGTGCTGCAGAAGGGAT (SEQ ID NO: 5539)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0975	PE2	PE2	FANCF	FANCF	+9 C to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATAGCTTTTCCGAGCTTCT (SEQ ID NO: 5540)	AGAAGCTCGGAAAAGCTATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5541)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0976	PE2	PE2	FANCF	FANCF	+9-10 CG to AT		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATATCTTTTCCGAGCTTCT (SEQ ID NO: 5542)	AGAAGCTCGGAAAAGATATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5543)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0977	PE2	PE2	FANCF	FANCF	+8-10 TCG to GAT		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGAGATCTTTTCCGAGCTTCT (SEQ ID NO: 5544)	AGAAGCTCGGAAAAGATCTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5545)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0978	PE2	PE2	FANCF	FANCF	+8-11 TCGC to GATA		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGAGATATTTTCCGAGCTTCT (SEQ ID NO: 5546)	AGAAGCTCGGAAAATATCTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5547)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0979	PE2	PE2	FANCF	FANCF	+7-11 ATCGC to CGATA		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGCGATATTTTCCGAGCTTCT (SEQ ID NO: 5548)	AGAAGCTCGGAAAATATCGCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5549)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0980	PE2	PE4	FANCF	FANCF	+3 C to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGTGGATCGCTTTTCCGAGC (SEQ ID NO: 5530)	GCTCGGAAAAGCGATCCACGTGCTGCAGAAGGGAT (SEQ ID NO: 5531)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0981	PE2	PE4	FANCF	FANCF	+3-4 CT to GA		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGAGGATCGCTTTTCCGAGC (SEQ ID NO: 5532)	GCTCGGAAAAGCGATCCTCGTGCTGCAGAAGGGAT (SEQ ID NO: 5533)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0982	PE2	PE4	FANCF	FANCF	+3-5 CTG to GAA		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGAAGATCGCTTTTCCGAGC (SEQ ID NO: 5534)	GCTCGGAAAAGCGATCTTCGTGCTGCAGAAGGGAT (SEQ ID NO: 5535)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0983	PE2	PE4	FANCF	FANCF	+3-6 CTGG to GAAT		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGAATATCGCTTTTCCGAGC (SEQ ID NO: 5536)	GCTCGGAAAAGCGATATTCGTGCTGCAGAAGGGAT (SEQ ID NO: 5537)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0984	PE2	PE4	FANCF	FANCF	+3-7 CTGGA to GAATC		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGAATCTCGCTTTTCCGAGC (SEQ ID NO: 5538)	GCTCGGAAAAGCGAGATTCGTGCTGCAGAAGGGAT (SEQ ID NO: 5539)	0	0	0	0	
44D,44E,50I	HEK293T	Chen2021_0985	PE2	PE4	FANCF	FANCF	+9 C to A		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATAGCTTTTCCGAGCTTCT (SEQ ID NO: 5540)	AGAAGCTCGGAAAAGCTATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5541)	0	0	0	0	

44D,44E,50I	HEK293T	Chen2021_0986	PE2	PE4	FANCF	FANCF	+9–10 CG to AT		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATATCTTTTCCGAGCTTCT (SEQ ID NO: 5542)	AGAAGCTCGGAAAAGATATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5543)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0987	PE2	PE4	FANCF	FANCF	+8–10 TCG to GAT		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGAGATCTTTTCCGAGCTTCT (SEQ ID NO: 5544)	AGAAGCTCGGAAAAGATCTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5545)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0988	PE2	PE4	FANCF	FANCF	+8–11 TCGC to GATA		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGAGATATTTTCCGAGCTTCT (SEQ ID NO: 5546)	AGAAGCTCGGAAAATATCTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5547)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0989	PE2	PE4	FANCF	FANCF	+7–11 ATCGC to CGATA		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGCGATATTTTCCGAGCTTCT (SEQ ID NO: 5548)	AGAAGCTCGGAAAATATCGCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5549)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0990	PE2	PE2	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAAA (SEQ ID NO: 5550)	TTTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5551)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0991	PE2	PE2	HEK3	HEK3	+1–2 TG to AT		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATATGGCAGAGGAAA (SEQ ID NO: 5552)	TTTCCTCTGCCATATCGTGCTCAGTCTG (SEQ ID NO: 5553)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0992	PE2	PE2	HEK3	HEK3	+1–3 TGA to ATC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATCTGGCAGAGGAAA (SEQ ID NO: 5554)	TTTCCTCTGCCAGATCGTGCTCAGTCTG (SEQ ID NO: 5555)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0993	PE2	PE2	HEK3	HEK3	+1–4 TGAT to ATCC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATCCGGCAGAGGAAA (SEQ ID NO: 5556)	TTTCCTCTGCCGGATCGTGCTCAGTCTG (SEQ ID NO: 5557)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0994	PE2	PE2	HEK3	HEK3	+1–5 TGATG to ATCCA		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATCCAGCAGAGGAAA (SEQ ID NO: 5558)	TTTCCTCTGCTGGATCGTGCTCAGTCTG (SEQ ID NO: 5559)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0995	PE2	PE2	HEK3	HEK3	+8 A to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGCTGA (SEQ ID NO: 5560)	TCAGCCATCACGTGCTCAGTCTG (SEQ ID NO: 5561)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0996	PE2	PE2	HEK3	HEK3	+8–9 AG to TC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGCTCAGGAAAGGAAGCCCT (SEQ ID NO: 5562)	AGGGCTTCCTTTCCTGAGCCATCACGTGCTCAGTCTG (SEQ ID NO: 5563)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0997	PE2	PE2	HEK3	HEK3	+8–10 AGA to TCC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGCTCCGGAAAGGAAGCCCT (SEQ ID NO: 5564)	AGGGCTTCCTTTCCGGAGCCATCACGTGCTCAGTCTG (SEQ ID NO: 5565)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0998	PE2	PE2	HEK3	HEK3	+7–10 CAGA to TTCC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTTCCGGAAAGGAAGCCCT (SEQ ID NO: 5566)	AGGGCTTCCTTTCCGGAACCATCACGTGCTCAGTCTG (SEQ ID NO: 5567)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_0999	PE2	PE2	HEK3	HEK3	+7–11 CAGAG to TTCCC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTTCCCGAAAGGAAGCCCT (SEQ ID NO: 5568)	AGGGCTTCCTTTCGGGAACCATCACGTGCTCAGTCTG (SEQ ID NO: 5569)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1000	PE2	PE4	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAAA (SEQ ID NO: 5550)	TTTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5551)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1001	PE2	PE4	HEK3	HEK3	+1–2 TG to AT		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATATGGCAGAGGAAA (SEQ ID NO: 5552)	TTTCCTCTGCCATATCGTGCTCAGTCTG (SEQ ID NO: 5553)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1002	PE2	PE4	HEK3	HEK3	+1–3 TGA to ATC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATCTGGCAGAGGAAA (SEQ ID NO: 5554)	TTTCCTCTGCCAGATCGTGCTCAGTCTG (SEQ ID NO: 5555)	0	0	0	0

44D,44E,50I	HEK293T	Chen2021_1003	PE2	PE4	HEK3	HEK3	+1–4 TGAT to ATCC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATCCGGCAGAGGAAA (SEQ ID NO: 5556)	TTTCCTCTGCCGGATCGTGCTCAGTCTG (SEQ ID NO: 5557)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1004	PE2	PE4	HEK3	HEK3	+1–5 TGATG to ATCCA		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	ATCCAGCAGAGGAAA (SEQ ID NO: 5558)	TTTCCTCTGCTGGATCGTGCTCAGTCTG (SEQ ID NO: 5559)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1005	PE2	PE4	HEK3	HEK3	+8 A to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGCTGA (SEQ ID NO: 5560)	TCAGCCATCACGTGCTCAGTCTG (SEQ ID NO: 5561)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1006	PE2	PE4	HEK3	HEK3	+8–9 AG to TC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGCTCAGGAAAGGAAGCCCT (SEQ ID NO: 5562)	AGGGCTTCCTTTCCTGAGCCATCACGTGCTCAGTCTG (SEQ ID NO: 5563)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1007	PE2	PE4	HEK3	HEK3	+8–10 AGA to TCC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGCTCCGGAAAGGAAGCCCT (SEQ ID NO: 5564)	AGGGCTTCCTTTCCGGAGCCATCACGTGCTCAGTCTG (SEQ ID NO: 5565)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1008	PE2	PE4	HEK3	HEK3	+7–10 CAGA to TTCC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTTCCGGAAAGGAAGCCCT (SEQ ID NO: 5566)	AGGGCTTCCTTTCCGGAACCATCACGTGCTCAGTCTG (SEQ ID NO: 5567)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1009	PE2	PE4	HEK3	HEK3	+7–11 CAGAG to TTCCC		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTTCCCGAAAGGAAGCCCT (SEQ ID NO: 5568)	AGGGCTTCCTTTCGGGAACCATCACGTGCTCAGTCTG (SEQ ID NO: 5569)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1010	PE2	PE2	RNF2	RNF2	+9 T to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGTGATCGTTGTAACT (SEQ ID NO: 5570)	AGTTACAACGATCACCTCAGGTAATGACTAAGATG (SEQ ID NO: 5571)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1011	PE2	PE2	RNF2	RNF2	+8–9 GT to AA		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGTAATCGTTGTAACT (SEQ ID NO: 5572)	AGTTACAACGATTACCTCAGGTAATGACTAAGATG (SEQ ID NO: 5573)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1012	PE2	PE2	RNF2	RNF2	+8–10 GTT to AAG		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGTAAGCGTTGTAACT (SEQ ID NO: 5574)	AGTTACAACGCTTACCTCAGGTAATGACTAAGATG (SEQ ID NO: 5575)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1013	PE2	PE2	RNF2	RNF2	+7–10 TGTT to CAAG		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCAAGCGTTGTAACT (SEQ ID NO: 5576)	AGTTACAACGCTTGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5577)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1014	PE2	PE2	RNF2	RNF2	+7–11 TGTTC to CAAGA		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCAAGAGTTGTAACT (SEQ ID NO: 5578)	AGTTACAACTCTTGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5579)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1015	PE2	PE4	RNF2	RNF2	+9 T to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGTGATCGTTGTAACT (SEQ ID NO: 5570)	AGTTACAACGATCACCTCAGGTAATGACTAAGATG (SEQ ID NO: 5571)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1016	PE2	PE4	RNF2	RNF2	+8–9 GT to AA		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGTAATCGTTGTAACT (SEQ ID NO: 5572)	AGTTACAACGATTACCTCAGGTAATGACTAAGATG (SEQ ID NO: 5573)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1017	PE2	PE4	RNF2	RNF2	+8–10 GTT to AAG		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGTAAGCGTTGTAACT (SEQ ID NO: 5574)	AGTTACAACGCTTACCTCAGGTAATGACTAAGATG (SEQ ID NO: 5575)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1018	PE2	PE4	RNF2	RNF2	+7–10 TGTT to CAAG		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCAAGCGTTGTAACT (SEQ ID NO: 5576)	AGTTACAACGCTTGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5577)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1019	PE2	PE4	RNF2	RNF2	+7–11 TGTTC to CAAGA		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCAAGAGTTGTAACT (SEQ ID NO: 5578)	AGTTACAACTCTTGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5579)	0	0	0	0

44D,44E,50I	HEK293T	Chen2021_1020	PE2	PE2	VEGFA	VEGFA	+9 C to G		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCTGCAGATGGCACATTGT (SEQ ID NO: 5580)	ACAATGTGCCATCTGCAGCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5581)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1021	PE2	PE2	VEGFA	VEGFA	+9–10 CC to GT		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCTGTAGATGGCACATTGT (SEQ ID NO: 5582)	ACAATGTGCCATCTACAGCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5583)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1022	PE2	PE2	VEGFA	VEGFA	+8–10 TCC to GGT		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCGGTAGATGGCACATTGT (SEQ ID NO: 5584)	ACAATGTGCCATCTACCGCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5585)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1023	PE2	PE2	VEGFA	VEGFA	+8–11 TCCA to GGTT		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCGGTTGATGGCACATTGT (SEQ ID NO: 5586)	ACAATGTGCCATCAACCGCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5587)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1024	PE2	PE2	VEGFA	VEGFA	+7–11 CTCCA to AGGTT		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGAGGTTGATGGCACATTGT (SEQ ID NO: 5588)	ACAATGTGCCATCAACCTCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5589)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1025	PE2	PE4	VEGFA	VEGFA	+9 C to G		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCTGCAGATGGCACATTGT (SEQ ID NO: 5580)	ACAATGTGCCATCTGCAGCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5581)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1026	PE2	PE4	VEGFA	VEGFA	+9–10 CC to GT		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCTGTAGATGGCACATTGT (SEQ ID NO: 5582)	ACAATGTGCCATCTACAGCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5583)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1027	PE2	PE4	VEGFA	VEGFA	+8–10 TCC to GGT		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCGGTAGATGGCACATTGT (SEQ ID NO: 5584)	ACAATGTGCCATCTACCGCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5585)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1028	PE2	PE4	VEGFA	VEGFA	+8–11 TCCA to GGTT		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGCGGTTGATGGCACATTGT (SEQ ID NO: 5586)	ACAATGTGCCATCAACCGCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5587)	0	0	0	0
44D,44E,50I	HEK293T	Chen2021_1029	PE2	PE4	VEGFA	VEGFA	+7–11 CTCCA to AGGTT		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGAGGTTGATGGCACATTGT (SEQ ID NO: 5588)	ACAATGTGCCATCAACCTCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5589)	0	0	0	0
44G,50J	HeLa	Chen2021_1030	PE2–P2A–BSD	PE2	DNMT1	DNMT1	+1 A to G		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	GCAGGGGTGAC (SEQ ID NO: 5408)	GTCACCCCTGCTTCTGGCACCAGG (SEQ ID NO: 5409)	0	0	0	0
44F,44G	HeLa	Chen2021_1031	PE2–P2A–BSD	PE2	DNMT1	DNMT1	+6 G to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	0	0
44G,50J	HeLa	Chen2021_1032	PE2–P2A–BSD	PE2	DNMT1	DNMT1	+8 T to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGAGACGGGAGGGC (SEQ ID NO: 5424)	GCCCTCCCGTCTCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5425)	0	0	0	0
44G,50J	HeLa	Chen2021_1033	PE2–P2A–BSD	PE2	EMX1	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
44G,50J	HeLa	Chen2021_1034	PE2–P2A–BSD	PE2	EMX1	EMX1	+5 G to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	0	0
44G,50J	HeLa	Chen2021_1035	PE2–P2A–BSD	PE2	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
44F,44G	HeLa	Chen2021_1036	PE2–P2A–BSD	PE2	FANCF	FANCF	+3 C to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGTGGATCGCTTTTCC (SEQ ID NO: 5310)	GGAAAAGCGATCCACGTGCTGCAGAAGGGAT (SEQ ID NO: 5311)	0	0	0	0

44F,44G	HeLa	Chen2021_1037	PE2-P2A-BSD	PE2	FANCF	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0
44G,50J	HeLa	Chen2021_1038	PE2-P2A-BSD	PE2	FANCF	FANCF	+4 T to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGGGATCGCTTTTCC (SEQ ID NO: 5314)	GGAAAAGCGATCCCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5315)	0	0	0	0
44F,44G	HeLa	Chen2021_1039	PE2-P2A-BSD	PE2	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
44G,50J	HeLa	Chen2021_1040	PE2-P2A-BSD	PE2	HEK3	HEK3	+1 T to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5324)	TCTGCCATCGCGTGCTCAGTCTG (SEQ ID NO: 5325)	0	0	0	0
44G,50J	HeLa	Chen2021_1041	PE2-P2A-BSD	PE2	HEK3	HEK3	+1 T to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGATGGCAGA (SEQ ID NO: 5326)	TCTGCCATCCCGTGCTCAGTCTG (SEQ ID NO: 5327)	0	0	0	0
44G,50J	HeLa	Chen2021_1042	PE2-P2A-BSD	PE2	HEK3	HEK3	+3 A to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGCTGGCAGA (SEQ ID NO: 5328)	TCTGCCAGCACGTGCTCAGTCTG (SEQ ID NO: 5329)	0	0	0	0
44G,50J	HeLa	Chen2021_1043	PE2-P2A-BSD	PE2	HEK3	HEK3	+3 A to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGTGGCAGA (SEQ ID NO: 5330)	TCTGCCACCACGTGCTCAGTCTG (SEQ ID NO: 5331)	0	0	0	0
44G,50J	HeLa	Chen2021_1044	PE2-P2A-BSD	PE2	HEK3	HEK3	+3 A to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5332)	TCTGCCAACACGTGCTCAGTCTG (SEQ ID NO: 5333)	0	0	0	0
44G,50J	HeLa	Chen2021_1045	PE2-P2A-BSD	PE2	HEK3	HEK3	+5 G to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATAGCAGA (SEQ ID NO: 5334)	TCTGCTATCACGTGCTCAGTCTG (SEQ ID NO: 5335)	0	0	0	0
44G,50J	HeLa	Chen2021_1046	PE2-P2A-BSD	PE2	HEK3	HEK3	+5 G to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGCAGA (SEQ ID NO: 5336)	TCTGCGATCACGTGCTCAGTCTG (SEQ ID NO: 5337)	0	0	0	0
44G,50J	HeLa	Chen2021_1047	PE2-P2A-BSD	PE2	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0
44G,50J	HeLa	Chen2021_1048	PE2-P2A-BSD	PE2	HEK3	HEK3	+7 C to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGAAGA (SEQ ID NO: 5338)	TCTTCCATCACGTGCTCAGTCTG (SEQ ID NO: 5339)	0	0	0	0
44G,50J	HeLa	Chen2021_1049	PE2-P2A-BSD	PE2	HEK3	HEK3	+7 C to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGGAGA (SEQ ID NO: 5340)	TCTCCCATCACGTGCTCAGTCTG (SEQ ID NO: 5341)	0	0	0	0
44G,50J	HeLa	Chen2021_1050	PE2-P2A-BSD	PE2	HEK3	HEK3	+7 C to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTAGA (SEQ ID NO: 5342)	TCTACCATCACGTGCTCAGTCTG (SEQ ID NO: 5343)	0	0	0	0
44G,50J	HeLa	Chen2021_1051	PE2-P2A-BSD	PE2	RNF2	RNF2	+4 A to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGTGGTGTTCGTT (SEQ ID NO: 5378)	AACGAACACCACAGGTAATGACTAAGATG (SEQ ID NO: 5379)	0	0	0	0
44F,44G	HeLa	Chen2021_1052	PE2-P2A-BSD	PE2	RNF2	RNF2	+6 G to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	0	0
44G,50J	HeLa	Chen2021_1053	PE2-P2A-BSD	PE2	RNF2	RNF2	+7 T to C		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCGTTCGTT (SEQ ID NO: 5384)	AACGAACGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5385)	0	0	0	0

44F,44G	HeLa	Chen2021_1054	PE2–P2A–BSD	PE2	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
44G,50J	HeLa	Chen2021_1055	PE2–P2A–BSD	PE2	RUNX1	RUNX1	+4 T to A		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGAAGGCTTCAGACA (SEQ ID NO: 5398)	TGTCTGAAGCCTTCGCTTCCTCCTGAAAAT (SEQ ID NO: 5399)	0	0	0	0
44F,44G	HeLa	Chen2021_1056	PE2–P2A–BSD	PE2	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
44G,50J	HeLa	Chen2021_1057	PE2–P2A–BSD	PE2	VEGFA	VEGFA	+1 T to G		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	GGAGGGCTCCAGATGGCACATT (SEQ ID NO: 5432)	AATGTGCCATCTGGAGCCCTCCTCTGGCCTGCAGA (SEQ ID NO: 5433)	0	0	0	0
44F,44G	HeLa	Chen2021_1058	PE2–P2A–BSD	PE2	VEGFA	VEGFA	+7 C to A		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	0	0
44G,50J	HeLa	Chen2021_1059	PE2–P2A–BSD	PE2	VEGFA	VEGFA	+7 C to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGTTCCAGATGGCACATT (SEQ ID NO: 5446)	AATGTGCCATCTGGAACCCTCATCTGGCCTGCAGA (SEQ ID NO: 5447)	0	0	0	0
44G,50J	HeLa	Chen2021_1060	PE2–P2A–BSD	PE4	DNMT1	DNMT1	+1 A to G		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	GCAGGGGTGAC (SEQ ID NO: 5408)	GTCACCCCTGCTTCTGGCACCAGG (SEQ ID NO: 5409)	0	0	0	0
44F,44G	HeLa	Chen2021_1061	PE2–P2A–BSD	PE4	DNMT1	DNMT1	+6 G to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	0	0
44G,50J	HeLa	Chen2021_1062	PE2–P2A–BSD	PE4	DNMT1	DNMT1	+8 T to A		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGAGACGGGAGGGC (SEQ ID NO: 5424)	GCCCTCCCGTCTCCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5425)	0	0	0	0
44G,50J	HeLa	Chen2021_1063	PE2–P2A–BSD	PE4	EMX1	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
44G,50J	HeLa	Chen2021_1064	PE2–P2A–BSD	PE4	EMX1	EMX1	+5 G to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	0	0
44G,50J	HeLa	Chen2021_1065	PE2–P2A–BSD	PE4	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
44F,44G	HeLa	Chen2021_1066	PE2–P2A–BSD	PE4	FANCF	FANCF	+3 C to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACGTGGATCGCTTTTCC (SEQ ID NO: 5310)	GGAAAAGCGATCCACGTGCTGCAGAAGGGAT (SEQ ID NO: 5311)	0	0	0	0
44F,44G	HeLa	Chen2021_1067	PE2–P2A–BSD	PE4	FANCF	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0
44G,50J	HeLa	Chen2021_1068	PE2–P2A–BSD	PE4	FANCF	FANCF	+4 T to G		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGGGATCGCTTTTCC (SEQ ID NO: 5314)	GGAAAAGCGATCCCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5315)	0	0	0	0
44F,44G	HeLa	Chen2021_1069	PE2–P2A–BSD	PE4	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
44G,50J	HeLa	Chen2021_1070	PE2–P2A–BSD	PE4	HEK3	HEK3	+1 T to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5324)	TCTGCCATCGCGTGCTCAGTCTG (SEQ ID NO: 5325)	0	0	0	0

44G,50J	HeLa	Chen2021_1071	PE2--P2A--BSD	PE4	HEK3	HEK3	+1 T to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGATGGCAGA (SEQ ID NO: 5326)	TCTGCCATCCCGTGCTCAGTCTG (SEQ ID NO: 5327)	0	0	0
44G,50J	HeLa	Chen2021_1072	PE2--P2A--BSD	PE4	HEK3	HEK3	+3 A to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGCTGGCAGA (SEQ ID NO: 5328)	TCTGCCAGCACGTGCTCAGTCTG (SEQ ID NO: 5329)	0	0	0
44G,50J	HeLa	Chen2021_1073	PE2--P2A--BSD	PE4	HEK3	HEK3	+3 A to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGTGGCAGA (SEQ ID NO: 5330)	TCTGCCACCACGTGCTCAGTCTG (SEQ ID NO: 5331)	0	0	0
44G,50J	HeLa	Chen2021_1074	PE2--P2A--BSD	PE4	HEK3	HEK3	+3 A to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5332)	TCTGCCAACACGTGCTCAGTCTG (SEQ ID NO: 5333)	0	0	0
44G,50J	HeLa	Chen2021_1075	PE2--P2A--BSD	PE4	HEK3	HEK3	+5 G to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATAGCAGA (SEQ ID NO: 5334)	TCTGCTATCACGTGCTCAGTCTG (SEQ ID NO: 5335)	0	0	0
44G,50J	HeLa	Chen2021_1076	PE2--P2A--BSD	PE4	HEK3	HEK3	+5 G to C		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGCAGA (SEQ ID NO: 5336)	TCTGCGATCACGTGCTCAGTCTG (SEQ ID NO: 5337)	0	0	0
44G,50J	HeLa	Chen2021_1077	PE2--P2A--BSD	PE4	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0
44G,50J	HeLa	Chen2021_1078	PE2--P2A--BSD	PE4	HEK3	HEK3	+7 C to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGAAGA (SEQ ID NO: 5338)	TCTTCCATCACGTGCTCAGTCTG (SEQ ID NO: 5339)	0	0	0
44G,50J	HeLa	Chen2021_1079	PE2--P2A--BSD	PE4	HEK3	HEK3	+7 C to G		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGGAGA (SEQ ID NO: 5340)	TCTCCCATCACGTGCTCAGTCTG (SEQ ID NO: 5341)	0	0	0
44G,50J	HeLa	Chen2021_1080	PE2--P2A--BSD	PE4	HEK3	HEK3	+7 C to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTAGA (SEQ ID NO: 5342)	TCTACCATCACGTGCTCAGTCTG (SEQ ID NO: 5343)	0	0	0
44G,50J	HeLa	Chen2021_1081	PE2--P2A--BSD	PE4	RNF2	RNF2	+4 A to T		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGTGGTGTTCGTT (SEQ ID NO: 5378)	AACGAACACCACAGGTAATGACTAAGATG (SEQ ID NO: 5379)	0	0	0
44F,44G	HeLa	Chen2021_1082	PE2--P2A--BSD	PE4	RNF2	RNF2	+6 G to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	0
44G,50J	HeLa	Chen2021_1083	PE2--P2A--BSD	PE4	RNF2	RNF2	+7 T to C		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCGTTCGTT (SEQ ID NO: 5384)	AACGAACGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5385)	0	0	0
44F,44G	HeLa	Chen2021_1084	PE2--P2A--BSD	PE4	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0
44G,50J	HeLa	Chen2021_1085	PE2--P2A--BSD	PE4	RUNX1	RUNX1	+4 T to A		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGAAGGCTTCAGACA (SEQ ID NO: 5398)	TGTCTGAAGCCTTCGCTTCCTCCTGAAAAT (SEQ ID NO: 5399)	0	0	0
44F,44G	HeLa	Chen2021_1086	PE2--P2A--BSD	PE4	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO. 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0
44G,50J	HeLa	Chen2021_1087	PE2--P2A--BSD	PE4	VEGFA	VEGFA	+1 T to G		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	GGAGGGCTCCAGATGGCACATT (SEQ ID NO: 5432)	AATGTGCCATCTGGAGCCCTCCTCTGGCCTGCAGA (SEQ ID NO: 5433)	0	0	0

44F,44G	HeLa	Chen2021_1088	PE2-P2A-SSD	PE4	VEGFA	VEGFA	+7 C to A		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGGCCAGAA (SEQ ID NO: 5266)	TGACGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCG ATCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	0	0	
44G,50J	HeLa	Chen2021_1089	PE2-P2A-SSD	PE4	VEGFA	VEGFA	+7 C to T		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5266)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGGCCAGAA (SEQ ID NO: 5266)	TGACGGTTCCAGAT GGCACATT (SEQ ID NO: 5446)	AATGTGCCATCTGGAAC CCTCATCTGGCCTGCAG A (SEQ ID NO: 5447)	0	0	0	0	
44G,50J	HeLa	Chen2021_1090	PE2-P2A-SSD	PE3	DNMT1	DNMT1	+1 A to G	+49	GATTCCGTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAAA (SEQ ID NO: 5262)	GCAGGGGTGAC CACCAGG (SEQ ID NO: 5438)	GTCACCCCTGGTTCTGG CACCAGG (SEQ ID NO: 5439)	0	0	GCCCTTCAG CTAAAATAAA TGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44F,44G	HeLa	Chen2021_1091	PE2-P2A-SSD	PE3	DNMT1	DNMT1	+5 G to C	+48	GATTCCGTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAAA (SEQ ID NO: 5262)	ACAGGGCGTGAC CACCAGG (SEQ ID NO: 5422)	GTCACGCGTTCTTCTGG CACCAGG (SEQ ID NO: 5423)	0	0	GCCCTTCAG CTAAAATAAA TGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44G,50J	HeLa	Chen2021_1092	PE2-P2A-SSD	PE3	DNMT1	DNMT1	+8 T to A	+46	GATTCCGTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAAA (SEQ ID NO: 5262)	ACAGGGGAGACGG GACGGGC (SEQ ID NO: 5424)	GCCGCTCCCGTCTCCCCT GTTTCTGGCACCAGG (SEQ ID NO: 5425)	0	0	GCCCTTCAG CTAAAATAAA TGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44G,50J	HeLa	Chen2021_1093	PE2-P2A-SSD	PE3	EMX1	EMX1	+2 A to C	+53	CAGTCCGGACCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GCAGGGCTCCCATC AC (SEQ ID NO: 5344)	GTTGATGGCAGCCCTGCTC TCTTCTGTCGTCGGA (SEQ ID NO: 5345)	0	0	GACATCGAT GTCCTCCCCC TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44G,50J	HeLa	Chen2021_1094	PE2-P2A-SSD	PE3	EMX1	EMX1	+5 G to A	+53	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGAGCTCCCATC AC (SEQ ID NO: 5246)	GTGATGGGAGCACTTCT CTTCTGTCGTCGGA (SEQ ID NO: 5250)	0	0	GACATCGAT GTCCTCCCCC TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44G,50J	HeLa	Chen2021_1095	PE2-P2A-SSD	PE3	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATC AC (SEQ ID NO: 5238)	GTGATGGGAGCACTTCT CTTCTGTCGTCGGA (SEQ ID NO: 5250)	0	0	GACATCGAT GTCCTCCCCC TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44F,44G	HeLa	Chen2021_1096	PE2-P2A-SSD	PE3	FANCF	FANCF	+3 C to G	+48	GGAATCCGTTC TGCACGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5256)	AGGTGGATCGCTTT TCC (SEQ ID NO: 5310)	GGAAAAGCGATCCACGT GCTCGCAGAAGGGGAT (SEQ ID NO: 5311)	0	0	GGGGGTCCCA GGTGCTGAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44F,44G	HeLa	Chen2021_1097	PE2-P2A-SSD	PE3	FANCF	FANCF	+3 C to T	+48	GGAATCCGTTC TGCACGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5256)	ACTTGGATCGCTTT TCC (SEQ ID NO: 5312)	GGAAAAGCGATCCACGT GCTCGCAGAAGGGGAT (SEQ ID NO: 5313)	0	0	GGGGGTCCCA GGTGCTGAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44G,50J	HeLa	Chen2021_1098	PE2-P2A-SSD	PE3	FANCF	FANCF	+4 T to G	+48	CTGAATCCGTTC TGCACGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5256)	ACCGGGGATCGCTTT TCC (SEQ ID NO: 5314)	GGAAAAGCGATCCACGT GCTCGCAGAAGGGGAT (SEQ ID NO: 5315)	0	0	GGGGGTCCCA GGTGCTGAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44F,44G	HeLa	Chen2021_1099	PE2-P2A-SSD	PE3	HEK3	HEK3	+1 T to A	+60	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGAGCTGAGCACG (SEQ ID NO: 5232)	AGATCGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAG ACCGGGGATCGCTTT TCC (SEQ ID NO: 5248)	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44G,50J	HeLa	Chen2021_1100	PE2-P2A-SSD	PE3	HEK3	HEK3	+1 T to C	+60	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGAGCTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5324)	TCTGCCATCGCGTGCTCAG AGTCGTG (SEQ ID NO: 5325)	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	
44G,50J	HeLa	Chen2021_1101	PE2-P2A-SSD	PE3	HEK3	HEK3	+1 T to G	+60	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGAGCTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5326)	TCTGCCATCCCGTGCTCAG GTCGTG (SEQ ID NO: 5327)	0	0	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTGC (SEQ ID NO: 5237)	

44G,50J	HeLa	Chen2021_1102	PE2–P2A–BSD	PE3	HEK3	HEK3	+3 A to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGCTGGCAGA (SEQ ID NO: 5328)	TCTGCCAGCACGTGCTCAGTCTG (SEQ ID NO: 5329)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1103	PE2–P2A–BSD	PE3	HEK3	HEK3	+3 A to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGTGGCAGA (SEQ ID NO: 5330)	TCTGCCACCACGTGCTCAGTCTG (SEQ ID NO: 5331)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1104	PE2–P2A–BSD	PE3	HEK3	HEK3	+3 A to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5332)	TCTGCCAACACGTGCTCAGTCTG (SEQ ID NO: 5333)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1105	PE2–P2A–BSD	PE3	HEK3	HEK3	+5 G to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATAGCAGA (SEQ ID NO: 5334)	TCTGCTATCACGTGCTCAGTCTG (SEQ ID NO: 5335)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1106	PE2–P2A–BSD	PE3	HEK3	HEK3	+5 G to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGCAGA (SEQ ID NO: 5336)	TCTGCGATCACGTGCTCAGTCTG (SEQ ID NO: 5337)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1107	PE2–P2A–BSD	PE3	HEK3	HEK3	+5 G to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1108	PE2–P2A–BSD	PE3	HEK3	HEK3	+7 C to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGAAGA (SEQ ID NO: 5338)	TCTTCCATCACGTGCTCAGTCTG (SEQ ID NO: 5339)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1109	PE2–P2A–BSD	PE3	HEK3	HEK3	+7 C to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGGAGA (SEQ ID NO: 5340)	TCTCCCATCACGTGCTCAGTCTG (SEQ ID NO: 5341)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1110	PE2–P2A–BSD	PE3	HEK3	HEK3	+7 C to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTAGA (SEQ ID NO: 5342)	TCTACCATCACGTGCTCAGTCTG (SEQ ID NO: 5343)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1111	PE2–P2A–BSD	PE3	RNF2	RNF2	+4 A to T	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGTGGTGTTCGTT (SEQ ID NO: 5378)	AACGAACACCACAGGTAATGACTAAGATG (SEQ ID NO: 5379)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	HeLa	Chen2021_1112	PE2–P2A–BSD	PE3	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1113	PE2–P2A–BSD	PE3	RNF2	RNF2	+7 T to C	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCGTTCGTT (SEQ ID NO: 5384)	AACGAACGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5385)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	HeLa	Chen2021_1114	PE2–P2A–BSD	PE3	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44G.50J	HeLa	Chen2021_1115	PE2-P2A-BSD	PE3	RUNX1	RUNX1	+4 T to A	+38	GCATTTTCAGG AGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGAGGAGGA AG (SEQ ID NO: 5243)	CGAAGGGTTCAGAC A (SEQ ID NO: 5396)	TGTCTGAAGGCTTCGGCTT CCTCGTCGAAAAT (SEQ ID NO: 5339)	0	0	GATGAAGCGA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44F.44G	HeLa	Chen2021_1116	PE2-P2A-BSD	PE3	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGG AGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGAGGAGGA AG (SEQ ID NO: 5243)	CGATTGGTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGGCAATCGGCTT CCTCGTCGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44G.50J	HeLa	Chen2021_1117	PE2-P2A-BSD	PE3	VEGFA	VEGFA	+1 T to G	+57	GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGGCCAGA (SEQ ID NO: 5266)	GGAGGGGTCCAGA TGGCACATT (SEQ ID NO: 5432)	AATGTGCCATGTGGAGC A (SEQ ID NO: 5433)	0	0	GATGTACAG AGAGCCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44F.44G	HeLa	Chen2021_1118	PE2-P2A-BSD	PE3	VEGFA	VEGFA	+7 C to A	-57	GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGGCCAGA (SEQ ID NO: 5266)	TGAGGGGTCCAGA TGGCACATT (SEQ ID NO: 5444)	AATGTGCCATGTGGAGC A (SEQ ID NO: 5445)	0	0	GATGTACAG AGAGCCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44G.50J	HeLa	Chen2021_1119	PE2-P2A-BSD	PE3	VEGFA	VEGFA	+7 C to T	+57	GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGGCCAGA (SEQ ID NO: 5266)	TGAGGGGTTCCAGAT GGCACATT (SEQ ID NO: 5446)	AATGTGCCATGTGGAAC A (SEQ ID NO: 5447)	0	0	GATGTACAG AGAGCCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44G.50J	HeLa	Chen2021_1120	PE2-P2A-BSD	PE5	DNMT1	DNMT1	+1 A to G	+49	GATTCCGTCGTG CCAGAAAGA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTTGGTGCCAGAAA (SEQ ID NO: 5262)	CCGAGGCGTGAC (SEQ ID NO: 5438)	GTCACCGCCTGGTTCTGG CACCAGG (SEQ ID NO: 5439)	0	0	GCCCTTCAG CTAAAATAAAT GG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44F.44G	HeLa	Chen2021_1121	PE2-P2A-BSD	PE5	DNMT1	DNMT1	+6 G to C	+49	GATTCCGTGGTG CCAGAAAGA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAAAG (SEQ ID NO: 5262)	ACAGGGCGTGAC (SEQ ID NO: 5422)	GTCAGCGCTGTTTCTGG CACCAGG (SEQ ID NO: 5423)	0	0	GCCCTTCAG CTAAAATAAAT GG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44G.50J	HeLa	Chen2021_1122	PE2-P2A-BSD	PE5	DNMT1	DNMT1	+6 T to A	+49	GATTCCGTGGTG CCAGAAAGA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAAAG (SEQ ID NO: 5262)	ACAGGGGGAGACGGG GAGGGGC (SEQ ID NO: 5424)	GCCCTCCCCGTCTCCCCT CTTTCTGGCACCAGG (SEQ ID NO: 5425)	0	0	GCCCTTCAG CTAAAATAAAT GG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44G.50J	HeLa	Chen2021_1123	PE2-P2A-BSD	PE5	EMX1	EMX1	+2 A to C	+53	GAGTCCGAGCAG GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGTCAGAAGAA (SEQ ID NO: 5239)	GCAGGGGTCCCCATC (SEQ ID NO: 5344)	GTTGATGGGAGCCCTTGCT TCTTCTGCTCCGGA (SEQ ID NO: 5345)	0	0	GACATCGAT GTCCTCCCC AT (SEQ ID NO: 5550)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44G.50J	HeLa	Chen2021_1124	PE2-P2A-BSD	PE5	EMX1	EMX1	+5 G to A	+53	GAGTCCGAGCAG GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCCATC (SEQ ID NO: 5246)	GTTGATGGGAGCTCTTCTTT CTTCTGCTCCGGA (SEQ ID NO: 5255)	0	0	GACATCGAT GTCCTCCCC AT (SEQ ID NO: 5550)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44G.50J	HeLa	Chen2021_1125	PE2-P2A-BSD	PE5	EMX1	EMX1	+5 G to T	+53	GAGTCCGCAGCA GAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCCATC (SEQ ID NO: 5236)	GTGATGGGAGCACTTCT TCTTCTGCTCCGGA (SEQ ID NO: 5240)	0	0	GACATCGAT GTCCTCCCC AT (SEQ ID NO: 5550)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44F.44G	HeLa	Chen2021_1126	PE2-P2A-BSD	PE5	FANCF	FANCF	+3 C to G	+48	GGAATCCGCTTC TGCACGACC (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5256)	ACTTGGATCGCTTT TCC (SEQ ID NO: 5316)	GGAAAAGCGATCCACGT GCTGCAGAAGGAT (SEQ ID NO: 5211)	0	0	GGGGTCCCA GGTGCTCAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
44F.44G	HeLa	Chen2021_1127	PE2-P2A-BSD	PE5	FANCF	FANCF	+3 C to T	+48	GGAATCCGCTTC TGCACGACC (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG (SEQ ID NO: 5256)	ACTTGGATCGCTTT TCC (SEQ ID NO: 5312)	GGAAAAGCGATCCACAAGT GCTGCAGAAGGAT (SEQ ID NO: 5213)	0	0	GGGGTCCCA GGTGCTCAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

44G,50J	HeLa	Chen2021_1128	PE2-P2A-BSD	PE5	FANCF	FANCF	+4 T to G	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCGGGATCGCTTTTCC (SEQ ID NO: 5314)	GGAAAAGCGATCCCGGTGCTGCAGAAGGGAT (SEQ ID NO: 5315)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	HeLa	Chen2021_1129	PE2-P2A-BSD	PE5	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1130	PE2-P2A-BSD	PE5	HEK3	HEK3	+1 T to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	CGATGGCAGA (SEQ ID NO: 5324)	TCTGCCATCGCGTGCTCAGTCTG (SEQ ID NO: 5325)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1131	PE2-P2A-BSD	PE5	HEK3	HEK3	+1 T to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	GGATGGCAGA (SEQ ID NO: 5326)	TCTGCCATCCCGTGCTCAGTCTG (SEQ ID NO: 5327)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1132	PE2-P2A-BSD	PE5	HEK3	HEK3	+3 A to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGCTGGCAGA (SEQ ID NO: 5328)	TCTGCCAGCACGTGCTCAGTCTG (SEQ ID NO: 5329)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1133	PE2-P2A-BSD	PE5	HEK3	HEK3	+3 A to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGGTGGCAGA (SEQ ID NO: 5330)	TCTGCCACCACGTGCTCAGTCTG (SEQ ID NO: 5331)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1134	PE2-P2A-BSD	PE5	HEK3	HEK3	+3 A to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGTTGGCAGA (SEQ ID NO: 5332)	TCTGCCAACACGTGCTCAGTCTG (SEQ ID NO: 5333)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1135	PE2-P2A-BSD	PE5	HEK3	HEK3	+5 G to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATAGCAGA (SEQ ID NO: 5334)	TCTGCTATCACGTGCTCAGTCTG (SEQ ID NO: 5335)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1136	PE2-P2A-BSD	PE5	HEK3	HEK3	+5 G to C	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATCGCAGA (SEQ ID NO: 5336)	TCTGCGATCACGTGCTCAGTCTG (SEQ ID NO: 5337)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1137	PE2-P2A-BSD	PE5	HEK3	HEK3	+5 G to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1138	PE2-P2A-BSD	PE5	HEK3	HEK3	+7 C to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGAAGA (SEQ ID NO: 5338)	TCTTCCATCACGTGCTCAGTCTG (SEQ ID NO: 5339)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1139	PE2-P2A-BSD	PE5	HEK3	HEK3	+7 C to G	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGGAGA (SEQ ID NO: 5340)	TCTCCCATCACGTGCTCAGTCTG (SEQ ID NO: 5341)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1140	PE2-P2A-BSD	PE5	HEK3	HEK3	+7 C to T	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGTAGA (SEQ ID NO: 5342)	TCTACCATCACGTGCTCAGTCTG (SEQ ID NO: 5343)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44G,50J	HeLa	Chen2021_1141	PE2–P2A–BSD	PE5	RNF2	RNF2	+4 A to T	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGTGGTGTTCGTT (SEQ ID NO: 5378)	AACGAACACCACAGGTAATGACTAAGATG (SEQ ID NO: 5379)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	HeLa	Chen2021_1142	PE2–P2A–BSD	PE5	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1143	PE2–P2A–BSD	PE5	RNF2	RNF2	+7 T to C	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGGCGTTCGTT (SEQ ID NO: 5384)	AACGAACGCCTCAGGTAATGACTAAGATG (SEQ ID NO: 5385)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	HeLa	Chen2021_1144	PE2–P2A–BSD	PE5	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1145	PE2–P2A–BSD	PE5	RUNX1	RUNX1	+4 T to A	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGAAGGCTTCAGACA (SEQ ID NO: 5398)	TGTCTGAAGCCTTCGCTTCCTCCTGAAAAT (SEQ ID NO: 5399)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	HeLa	Chen2021_1146	PE2–P2A–BSD	PE5	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1147	PE2–P2A–BSD	PE5	VEGFA	VEGFA	+1 T to G	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	GGAGGGCTCCAGATGGCACATT (SEQ ID NO: 5432)	AATGTGCCATCTGGAGCCCTCCTCTGGCCTGCAGA (SEQ ID NO: 5433)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	HeLa	Chen2021_1148	PE2–P2A–BSD	PE5	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G,50J	HeLa	Chen2021_1149	PE2–P2A–BSD	PE5	VEGFA	VEGFA	+7 C to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGTTCCAGATGGCACATT (SEQ ID NO: 5446)	AATGTGCCATCTGGAACCCTCATCTGGCCTGCAGA (SEQ ID NO: 5447)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G	HEK293T	Chen2021_1150	PE2	PE2	EMX1	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
44G	HEK293T	Chen2021_1151	PE2	PE2	EMX1	EMX1	+5 G to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	0	0
44G	HEK293T	Chen2021_1152	PE2	PE2	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
44G	HEK293T	Chen2021_1153	PE2	PE4	EMX1	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
44G	HEK293T	Chen2021_1154	PE2	PE4	EMX1	EMX1	+5 G to A		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	0	0
44G	HEK293T	Chen2021_1155	PE2	PE4	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0

44G	HEK293T	Chen2021_1156	PE2	PE3	EMX1	EMX1	+2 A to C	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G	HEK293T	Chen2021_1157	PE2	PE3	EMX1	EMX1	+5 G to A	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G	HEK293T	Chen2021_1158	PE2	PE3	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G	HEK293T	Chen2021_1159	PE2	PE5	EMX1	EMX1	+2 A to C	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G	HEK293T	Chen2021_1160	PE2	PE5	EMX1	EMX1	+5 G to A	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGAGCTCCCATCAC (SEQ ID NO: 5249)	GTGATGGGAGCTCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5250)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44G	HEK293T	Chen2021_1161	PE2	PE5	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	K562	Chen2021_1162	PE2	PE2	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
44F,44G	K562	Chen2021_1163	PE2	PE2	RNF2	RNF2	+6 G to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	0	0
44F,44G	K562	Chen2021_1164	PE2	PE2	FANCF	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0
44F,44G	K562	Chen2021_1165	PE2	PE2	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
44F,44G	K562	Chen2021_1166	PE2	PE2	DNMT1	DNMT1	+6 G to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	0	0
44F,44G	K562	Chen2021_1167	PE2	PE2	VEGFA	VEGFA	+7 C to A		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	0	0
44F,44G	K562	Chen2021_1168	PE2	PE4	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
44F,44G	K562	Chen2021_1169	PE2	PE4	RNF2	RNF2	+6 G to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	0	0
44F,44G	K562	Chen2021_1170	PE2	PE4	FANCF	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0

44F_44G	K562	Chen2021_1171	PE2	PE4	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGG AGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGAGGAGGA AG (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTAGAAAAT (SEQ ID NO: 5245)	0	0	0	0	
44F_44G	K562	Chen2021_1172	PE2	PE4	DNMT1	DNMT1	+6 G to C		GAATCCGTCGTTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTTGTGCCAGAA AG (SEQ ID NO: 5262)	ACACGCGTCAC CAGCAGG (SEQ ID NO: 5422)	GTCACGCTGTTTCTTG (SEQ ID NO: 5423)	0	0	0	0	
44F_44G	K562	Chen2021_1173	PE2	PE4	VEGFA	VEGFA	+7 C to A		GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGGCATCAGAT GGCACATT (SEQ ID NO: 5266)	TGACGGATCCACAT (SEQ ID NO: 5444)	AATGTGCCATCTGGA CTCATGTGGCTGCAGA (SEQ ID NO: 5445)	0	0	0	0	
44F_44G	K562	Chen2021_1174	PE2	PE3	HEK3	HEK3	+1 T to A	+49	GGGCCACAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACCG AGATGGCAGA (SEQ ID NO: 5232)	TCTGCCATCTCGTGCTCAA GTCTC (SEQ ID NO: 5246)	0	0	0	0	
44F_44G	K562	Chen2021_1175	PE2	PE3	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGT CATTACCTC (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA AC (SEQ ID NO: 5252)	CTGAGATTGTTCGTT TGAGTAAGATG (SEQ ID NO: 5383)	0	0	0	0	
44F_44G	K562	Chen2021_1176	PE2	PE3	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTC TGCACGACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG AC (SEQ ID NO: 5268)	ACTTGGATCGCTTT TCC (SEQ ID NO: 5312)	0	0	0	0	
44F_44G	K562	Chen2021_1177	PE2	PE3	RUNX1	RUNX1	+3 A to C	+38	CCATTTTCAGGG AGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTAGAAAAT (SEQ ID NO: 5245)	0	0	0	0	
44F_44G	K562	Chen2021_1178	PE2	PE3	DNMT1	DNMT1	+6 G to C	+49	GAATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA AG (SEQ ID NO: 5262)	CTCAGCCGTTGTAC CACCAGG (SEQ ID NO: 5423)	GTCACGCTGTTTCTTG (SEQ ID NO: 5423)	0	0	0	0	
44F_44G	K562	Chen2021_1179	PE2	PE3	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGGCATCAGAT GGCACATT (SEQ ID NO: 5266)	TGACGGATCCACAT (SEQ ID NO: 5444)	AATGTGCCATCTGGA CTCATGTGGCTGCAGA (SEQ ID NO: 5445)	0	0	0	0	
44F_44G	K562	Chen2021_1180	PE2	PE5	HEK3	HEK3	+1 T to A	+49	GGGCCACAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACCG AGATGGCAGA (SEQ ID NO: 5232)	TCTGCCATCTCGTGCTCAA GTCTG (SEQ ID NO: 5246)	0	0	0	0	
44F_44G	K562	Chen2021_1181	PE2	PE5	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGT CATTACCTC (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA AC (SEQ ID NO: 5252)	CTGAGATTGTTCGTT TGAGTAAGATG (SEQ ID NO: 5383)	0	0	0	0	
44F_44G	K562	Chen2021_1182	PE2	PE5	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTC TGCACGACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG AC (SEQ ID NO: 5256)	ACTTGGATCGCTTT TCC (SEQ ID NO: 5312)	0	0	0	0	
44F_44G	K562	Chen2021_1183	PE2	PE5	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGG AGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTAGAAAAT (SEQ ID NO: 5245)	0	0	0	0	
44F_44G	K562	Chen2021_1184	PE2	PE5	DNMT1	DNMT1	+6 G to C	+49	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA AG (SEQ ID NO: 5262)	ACAGGCGTCAC CAGCAGG (SEQ ID NO: 5422)	GTCACGCTGTTTCTTG (SEQ ID NO: 5423)	0	0	0	0	

44F,44G	K562	Chen2021_1165	PE2	PE5	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	U2OS	Chen2021_1166	PE2	PE2	FANCF	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0
44F,44G	U2OS	Chen2021_1167	PE2	PE2	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
44F,44G	U2OS	Chen2021_1168	PE2	PE2	DNMT1	DNMT1	+6 G to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	0	0
44F,44G	U2OS	Chen2021_1169	PE2	PE2	VEGFA	VEGFA	+7 C to A		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	0	0
44F,44G	U2OS	Chen2021_1190	PE2-P2A-MLH1dn	PE4	FANCF	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0
44F,44G	U2OS	Chen2021_1191	PE2-P2A-MLH1dn	PE4	RUNX1	RUNX1	+3 A to C		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	0	0
44F,44G	U2OS	Chen2021_1192	PE2-P2A-MLH1dn	PE4	DNMT1	DNMT1	+6 G to C		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	0	0
44F,44G	U2OS	Chen2021_1193	PE2-P2A-MLH1dn	PE4	VEGFA	VEGFA	+7 C to A		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	0	0
44F,44G	U2OS	Chen2021_1194	PE2	PE3	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	U2OS	Chen2021_1195	PE2	PE3	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	U2OS	Chen2021_1196	PE2	PE3	DNMT1	DNMT1	+6 G to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	U2OS	Chen2021_1197	PE2	PE3	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	U2OS	Chen2021_1198	PE2-P2A-MLH1dn	PE5	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	U2OS	Chen2021_1199	PE2-P2A-MLH1dn	PE5	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

44F,44G	U2OS	Chen2021_1200	PE2–P2A–MLH1dn	PE5	DNMT1	DNMT1	+6 G to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGGCACCAGG (SEQ ID NO: 5423)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
44F,44G	U2OS	Chen2021_1201	PE2–P2A–MLH1dn	PE5	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1202	untreated	untreated	HEK3								0	0	0	0	0
45B,45C,51A	HEK293T	Chen2021_1203	PE2	Control	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A	HEK293T	Chen2021_1204	PE2	MLH1dn	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A	HEK293T	Chen2021_1205	PE2	Control	HEK3	HEK3	non-editing	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGCAGA (SEQ ID NO: 5591)	TCTGCCATCACGTGCTCAGTCTG (SEQ ID NO: 5592)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A	HEK293T	Chen2021_1206	PE2	MLH1dn	HEK3	HEK3	non-editing	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATGGCAGA (SEQ ID NO: 5591)	TCTGCCATCACGTGCTCAGTCTG (SEQ ID NO: 5592)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1207	PE2–dRT	Control	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1208	PE2–dRT	MLH1dn	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1209	Cas9 H840A	Control	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1210	Cas9 H840A	MLH1dn	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1211	untreated	untreated	EMX1								0	0	0	0	0
45B,45C,51A	HEK293T	Chen2021_1212	PE2	Control	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A	HEK293T	Chen2021_1213	PE2	MLH1dn	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A	HEK293T	Chen2021_1214	PE2	Control	EMX1	EMX1	non-editing	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCTCCCATCAC (SEQ ID NO: 5593)	GTGATGGGAGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5594)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

45B,45C,51A	HEK293T	Chen2021_1215	PE2	MLH1dn	EMX1	EMX1	non-editing	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGGGCTCCCATCAC (SEQ ID NO: 5593)	GTGATGGGAGCCCTTCTTCTTCTGCTCGGA (SEQ ID NO: 5594)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1216	PE2--dRT	Control	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1217	PE2--dRT	MLH1dn	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1218	Cas9 H840A	Control	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1219	Cas9 H840A	MLH1dn	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1220	untreated	untreated	FANCF								0	0	0	0	0
45B,45C,51A	HEK293T	Chen2021_1221	PE2	Control	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A	HEK293T	Chen2021_1222	PE2	MLH1dn	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A	HEK293T	Chen2021_1223	PE2	Control	FANCF	FANCF	non-editing	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATCGCTTTTCC (SEQ ID NO: 5595)	GGAAAAGCGATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5596)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A	HEK293T	Chen2021_1224	PE2	MLH1dn	FANCF	FANCF	non-editing	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGATCGCTTTTCC (SEQ ID NO: 5595)	GGAAAAGCGATCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5596)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1225	PE2--dRT	Control	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1226	PE2--dRT	MLH1dn	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1227	Cas9 H840A	Control	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1228	Cas9 H840A	MLH1dn	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1229	untreated	untreated	DNMT1								0	0	0	0	0

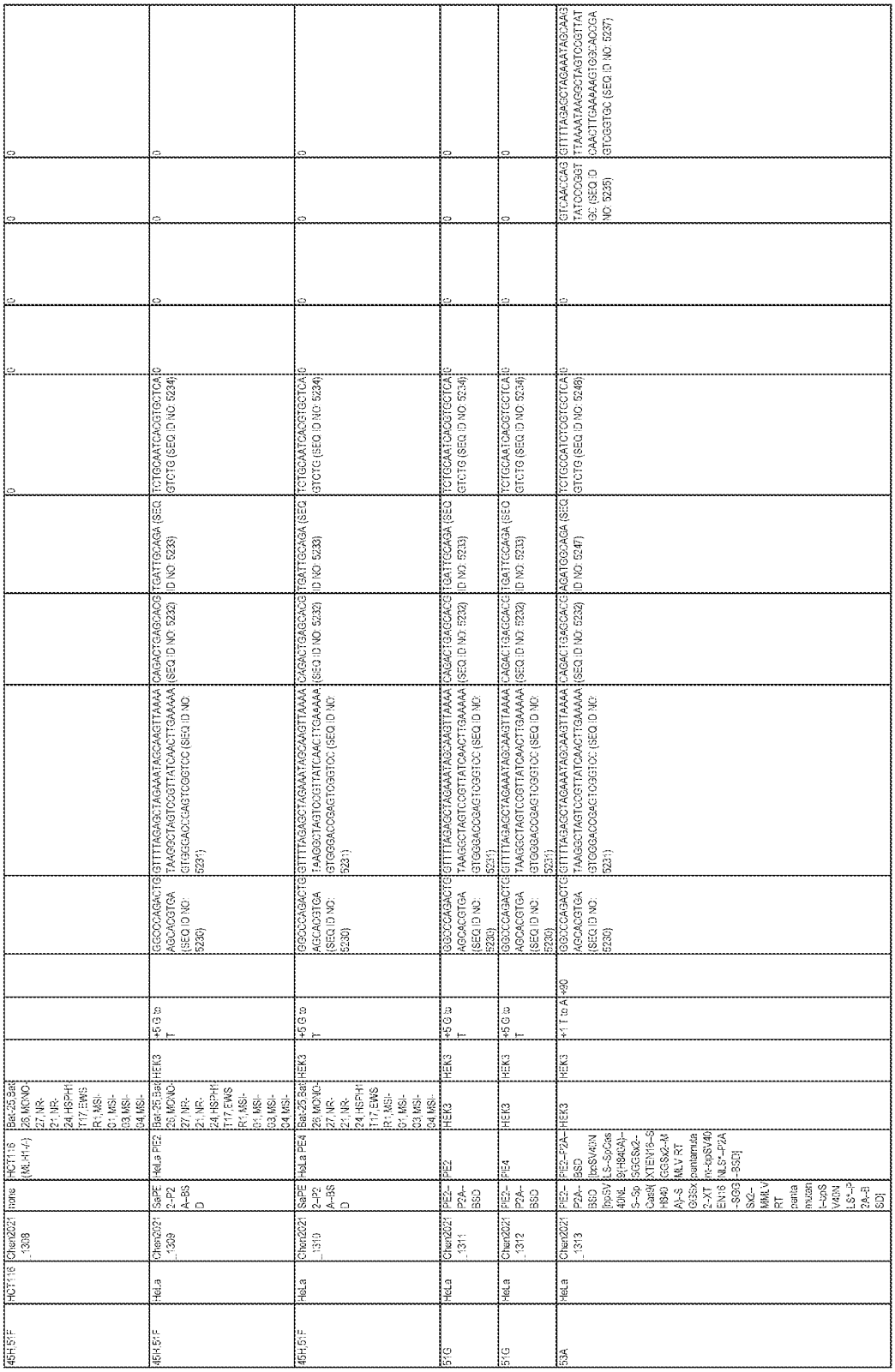
45B,45C,51A,51C	HEK293T	Chen2021_1230	PE2	Control	DNMT1	DNMT1	+1 A to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A,51C	HEK293T	Chen2021_1231	PE2	MLH1dn	DNMT1	DNMT1	+1 A to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A,51C	HEK293T	Chen2021_1232	PE2	Control	DNMT1	DNMT1	non-editing	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGTGAC (SEQ ID NO: 5597)	GTCACCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5598)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45B,45C,51A,51C	HEK293T	Chen2021_1233	PE2	MLH1dn	DNMT1	DNMT1	non-editing	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGGTGAC (SEQ ID NO: 5597)	GTCACCCCTGTTTCTGGCACCAGG (SEQ ID NO: 5598)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1234	PE2-dRT	Control	DNMT1	DNMT1	+1 A to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1235	PE2–dRT	MLH1dn	DNMT1	DNMT1	+1 A to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1236	Cas9 H840A	Control	DNMT1	DNMT1	+1 A to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
51A	HEK293T	Chen2021_1237	Cas9 H840A	MLH1dn	DNMT1	DNMT1	+1 A to C	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	CCAGGGGTGAC (SEQ ID NO: 5406)	GTCACCCCTGGTTCTGGCACCAGG (SEQ ID NO: 5407)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
45G	HEK293T	Chen2021_1238	PE2	PE2	HEK4	HEK4	+4 G to T		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCC (SEQ ID NO: 5600)	0	0	0	0
45G	HEK293T	Chen2021_1239	PE2	PE4	HEK4	HEK4	+4 G to T		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCC (SEQ ID NO: 5600)	0	0	0	0
51E	HEK293T	Chen2021_1240	PE2	PE2	HEK4	HEK4	+1 T to C		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCCGCCTCCAGCC (SEQ ID NO: 5602)	0	0	0	0
51E	HEK293T	Chen2021_1241	PE2	PE4	HEK4	HEK4	+1 T to C		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCCGCCTCCAGCC (SEQ ID NO: 5602)	0	0	0	0
45G	HEK293T	Chen2021_1242	PE2	PE2	HEK3 OT1	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
45G	HEK293T	Chen2021_1243	PE2	PE4	HEK3 OT1	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	0	0
51E	HEK293T	Chen2021_1244	PE2	PE2	HEK3 OT1	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCACGTGCTCAGTCTG (SEQ ID NO: 5234)	0	0	0	0

51E	HEK293T	Chen2021_1245	PE2	PE4	HEK3 OT1	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCCAATCAGCGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0	0	0
45G	HEK293T	Chen2021_1246	PE2	PE4	HEK3 OT2	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCAATCAGCGTGCTCA (SEQ ID NO: 5248)	0	0	0	0	0	0	0
45G	HEK293T	Chen2021_1247	PE2	PE4	HEK3 OT2	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCAATCAGCGTGCTCA (SEQ ID NO: 5248)	0	0	0	0	0	0	0
51E	HEK293T	Chen2021_1249	PE2	PE4	HEK3 OT2	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCCAATCAGCGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0	0	0
51E	HEK293T	Chen2021_1249	PE2	PE4	HEK3 OT2	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCCAATCAGCGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0	0	0
45G	HEK293T	Chen2021_1250	PE2	PE2	HEK3 OT3	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCA (SEQ ID NO: 5248)	0	0	0	0	0	0	0
45G	HEK293T	Chen2021_1251	PE2	PE4	HEK3 OT3	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCA (SEQ ID NO: 5248)	0	0	0	0	0	0	0
51E	HEK293T	Chen2021_1252	PE2	PE2	HEK3 OT3	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCAGCGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0	0	0
51E	HEK293T	Chen2021_1253	PE2	PE4	HEK3 OT3	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCAGCGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0	0	0
45G	HEK293T	Chen2021_1254	PE2	PE2	HEK3 OT4	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCA (SEQ ID NO: 5248)	0	0	0	0	0	0	0
45G	HEK293T	Chen2021_1256	PE2	PE4	HEK3 OT4	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCA (SEQ ID NO: 5248)	0	0	0	0	0	0	0
51E	HEK293T	Chen2021_1258	PE2	PE4	HEK3 OT4	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCAGCGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0	0	0
51E	HEK293T	Chen2021_1257	PE2	PE4	HEK3 OT4	HEK3	+5 G to T	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATCAGCGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0	0	0
45G	HEK293T	Chen2021_1258	PE2	EMX1 OT1	HEK3	+2 A to C	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5239)	GCAGGGCTCCCATC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTCA (SEQ ID NO: 5345)	0	0	0	0	0	0	0
45G	HEK293T	Chen2021_1259	PE2	EMX1 OT1	HEK3	+2 A to C	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5239)	GCAGGGCTCCCATC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTCA (SEQ ID NO: 5345)	0	0	0	0	0	0	0
51E	HEK293T	Chen2021_1260	PE2	EMX1 OT1	HEK3	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5239)	GAAGTGCTCCCATC (SEQ ID NO: 5299)	GTGATGGGAGCCACTTCT (SEQ ID NO: 5240)	0	0	0	0	0	0	0
51E	HEK293T	Chen2021_1261	PE2	PE4	EMX1 OT1	HEK3	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5239)	GAAGTGCTCCCATC (SEQ ID NO: 5299)	GTGATGGGAGCCACTTCT (SEQ ID NO: 5240)	0	0	0	0	0	0	0

45G	HEK293T	Chen2021_1262	PE2	PE2	EMX1 OT2	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
45G	HEK293T	Chen2021_1263	PE2	PE4	EMX1 OT2	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
51E	HEK293T	Chen2021_1264	PE2	PE2	EMX1 OT2	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
51E	HEK293T	Chen2021_1265	PE2	PE4	EMX1 OT2	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
45G	HEK293T	Chen2021_1266	PE2	PE2	EMX1 OT3	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
45G	HEK293T	Chen2021_1267	PE2	PE4	EMX1 OT3	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
51E	HEK293T	Chen2021_1268	PE2	PE2	EMX1 OT3	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
51E	HEK293T	Chen2021_1269	PE2	PE4	EMX1 OT3	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
45G	HEK293T	Chen2021_1270	PE2	PE2	EMX1 OT4	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
45G	HEK293T	Chen2021_1271	PE2	PE4	EMX1 OT4	EMX1	+2 A to C		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GCAGGGCTCCCATCAC (SEQ ID NO: 5344)	GTGATGGGAGCCCTGCTTCTTCTGCTCGGA (SEQ ID NO: 5345)	0	0	0	0
51E	HEK293T	Chen2021_1272	PE2	PE2	EMX1 OT4	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
51E	HEK293T	Chen2021_1273	PE2	PE4	EMX1 OT4	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
45G	HEK293T	Chen2021_1274	PE2	PE2	FANCF OT1	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0
45G	HEK293T	Chen2021_1275	PE2	PE4	FANCF OT1	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0
51E	HEK293T	Chen2021_1276	PE2	PE2	FANCF OT1	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)	0	0	0	0
51E	HEK293T	Chen2021_1277	PE2	PE4	FANCF OT1	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)	0	0	0	0
45G	HEK293T	Chen2021_1278	PE2	PE2	FANCF OT2	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	0	0

45G	HEK293T	Chen2021_1279	PE2	PE4	FANCF OT2	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)		0	0	0	0
51E	HEK293T	Chen2021_1280	PE2	PE2	FANCF OT2	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)		0	0	0	0
51E	HEK293T	Chen2021_1281	PE2	PE4	FANCF OT2	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)		0	0	0	0
45G	HEK293T	Chen2021_1282	PE2	PE2	FANCF OT3	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)		0	0	0	0
45G	HEK293T	Chen2021_1283	PE2	PE4	FANCF OT3	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)		0	0	0	0
51E	HEK293T	Chen2021_1284	PE2	PE2	FANCF OT3	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)		0	0	0	0
51E	HEK293T	Chen2021_1285	PE2	PE4	FANCF OT3	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)		0	0	0	0
45G	HEK293T	Chen2021_1286	PE2	PE2	FANCF OT4	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)		0	0	0	0
45G	HEK293T	Chen2021_1287	PE2	PE4	FANCF OT4	FANCF	+3 C to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)		0	0	0	0
51E	HEK293T	Chen2021_1288	PE2	PE2	FANCF OT4	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)		0	0	0	0
51E	HEK293T	Chen2021_1289	PE2	PE4	FANCF OT4	FANCF	+8 T to C		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACCTGGACCGCTTTTCC (SEQ ID NO: 5320)	GGAAAAGCGGTCCAGGTGCTGCAGAAGGGAT (SEQ ID NO: 5321)		0	0	0	0
45G	HEK293T	Chen2021_1290	PE2	PE2	HEK4 OT1	HEK4	+4 G to T		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCC (SEQ ID NO: 5600)		0	0	0	0
45G	HEK293T	Chen2021_1291	PE2	PE4	HEK4 OT1	HEK4	+4 G to T		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCC (SEQ ID NO: 5600)		0	0	0	0
51E	HEK293T	Chen2021_1292	PE2	PE2	HEK4 OT1	HEK4	+1 T to C		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCCGCCTCCAGCC (SEQ ID NO: 5602)		0	0	0	0
51E	HEK293T	Chen2021_1293	PE2	PE4	HEK4 OT1	HEK4	+1 T to C		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCCGCCTCCAGCC (SEQ ID NO: 5602)		0	0	0	0
45G	HEK293T	Chen2021_1294	PE2	PE2	HEK4 OT2	HEK4	+4 G to T		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCC (SEQ ID NO: 5600)		0	0	0	0
45G	HEK293T	Chen2021_1295	PE2	PE4	HEK4 OT2	HEK4	+4 G to T		GGCACTGCGGCTGGAGGTGG (SEQ ID NO: 5295)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCC (SEQ ID NO: 5600)		0	0	0	0

51E	HEK293T	Chen2021_1296	PE2	HEK4 OT2	HEK4	+1 T to C		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCGGCCTCCAG CC (SEQ ID NO: 5602)	0	0	0	0	0
51E	HEK293T	Chen2021_1297	PE2	HEK4 OT2	HEK4	+1 T to C		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCGGCCTCCAG CC (SEQ ID NO: 5602)	0	0	0	0	0
45G	HEK293T	Chen2021_1298	PE2	HEK4 OT3	HEK4	+4 G to T		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCCC (SEQ ID NO: 5600)	0	0	0	0	0
45G	HEK293T	Chen2021_1299	PE2	HEK4 OT3	HEK4	+4 G to T		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCCC (SEQ ID NO: 5600)	0	0	0	0	0
51E	HEK293T	Chen2021_1300	PE2	HEK4 OT3	HEK4	+1 T to C		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCGGCCTCCAG CC (SEQ ID NO: 5602)	0	0	0	0	0
51E	HEK293T	Chen2021_1301	PE2	HEK4 OT3	HEK4	+1 T to C		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCGGCCTCCAG CC (SEQ ID NO: 5602)	0	0	0	0	0
45G	HEK293T	Chen2021_1302	PE2	HEK4 OT4	HEK4	+4 G to T		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCCC (SEQ ID NO: 5600)	0	0	0	0	0
45G	HEK293T	Chen2021_1303	PE2	HEK4 OT4	HEK4	+4 G to T		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	TGGTGGTTAA (SEQ ID NO: 5599)	TTAACCACCACCTCCAGCCC (SEQ ID NO: 5600)	0	0	0	0	0
51E	HEK293T	Chen2021_1304	PE2	HEK4 OT4	HEK4	+1 T to C		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCGGCCTCCAG CC (SEQ ID NO: 5602)	0	0	0	0	0
51E	HEK293T	Chen2021_1305	PE2	HEK4 OT4	HEK4	+1 T to C		GGCACTGCGG CTGGAGGTGG (SEQ ID NO: 5296)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGCTGGAGG (SEQ ID NO: 5296)	CGGGGGTTAA (SEQ ID NO: 5601)	TTAACCCCGGCCTCCAG CC (SEQ ID NO: 5602)	0	0	0	0	0
45H-51E	HAP1 wild-type	Chen2021_1306	none	HAP1 wt	Bal:25 Bal:26 MONO 27 NR-21; NR-24 HSPH1 T17 EWSR1 MSI-01; MSI-03; MSI-04 MSI-								0	0	0	0	0
45H-51E	HAP1 ΔMSH2	Chen2021_1307	none	HAP1 ΔMSH2	Bal:25 Bal:26 MONO 27 NR-21; NR-24 HSPH1 T17 EWSR1 MSI-01; MSI-03; MSI-04 MSI-								0	0	0	0	0



45H45F																			
45H45F	HCT116	Chen2021 :1309	hsna	HCT116 (ATCC-CCL-247)	Bar25.Bar26.MONO27.NR-21.NR-24.HSPH1T17.EWSR1.MSI-01.MSI-03.MSI-04.MSI-	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATGACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
45H45F	HeLa	Chen2021 :1309	SaPE2-P2A-BSD	HeLa PE2	Bar25.Bar26.MONO27.NR-21.NR-24.HSPH1T17.EWSR1.MSI-01.MSI-03.MSI-04.MSI-	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATGACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
45H45F	HeLa	Chen2021 :1310	SaPE2-P2A-BSD	HeLa PE4	Bar25.Bar26.MONO27.NR-21.NR-24.HSPH1T17.EWSR1.MSI-01.MSI-03.MSI-04.MSI-	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATGACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
51G	HeLa	Chen2021 :1311	PE2-P2A-BSD	PE2	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATGACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
51G	HeLa	Chen2021 :1312	PE2-P2A-BSD	PE4	HEK3	HEK3	+5 G to T		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	TGATTGCAGA (SEQ ID NO: 5233)	TCTGCAATGACGTGCTCA (SEQ ID NO: 5234)	0	0	0	0	0
53A	HeLa	Chen2021 :1313	PE2-P2A-BSD [bpSV40 (ppSV40) - SV40 Late Poly(A) Signal] - [GGSGSG2-XTEN16-SGGS] - [MMLV RT (G33S) plantamute 2-XT mc-bpSV40 NLS*-P2A-BSD]	HEK3	+1 T to A	+80	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGCTCA (SEQ ID NO: 5248)	0	GTCAACCAG TATCCGGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5237)		

EGFA	HeLa	Chen2021_1314	pbSV40NLS-SpCas9-(HA6A)-SGGSx2-XTEN16-SGGSx2-M-MLV RT pentamutant GS pcon-pbSV40NLS*-myc-NLS-P2A-BSD	HEK3	+1 T to A #90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGTCATG (SEQ ID NO: 5248)			GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
EGFA	HeLa	Chen2021_1315	pbSV40NLS-SpCas9-(HA6A)-SGGSx2-XTEN16-SGGSx2-M-MLV RT pentamutant GS pcon-pbSV40NLS*-myc-NLS-P2A-BSD	HEK3	+1 T to A #90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGTCATG (SEQ ID NO: 5248)			GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1316	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–bpSV40NLS–P2A–BSD	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–bpSV40NLS–P2A–BSD	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0		0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1317	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0		0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1318	ppSV40NLS-SpCas9-2A-BSD	HEK3	+T to A+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGTCATG (SEQ ID NO: 5248)	G	GTCAACCAAGTATCCCGTGT (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	
53A	HeLa	Chen2021_1319	ppSV40NLS-SpCas9-2A-BSD	HEK3	+T to A+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGTCATG (SEQ ID NO: 5248)	G	GTCAACCAAGTATCCCGTGT (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	
53A	HeLa	Chen2021_1320	ppSV40NLS-SpCas9-2A-BSD	HEK3	+T to A+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGTCATG (SEQ ID NO: 5248)	G	GTCAACCAAGTATCCCGTGT (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	

53A	HeLa	Chen2021_1321	pbSV40NLS-SpCas9-(R221K N394K H840A)-GSGS-MLV RT pentamutant GS coon-opt SV40NLS-MMLV RT pentamutant GS coon-opt SV40NLS-P2A-BSD	HEK3	+1 T to A #40	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTGGCC (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGGCATCTCGTCGTCATG (SEQ ID NO: 5248)			GTCAACCAGTATCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1322	pbSV40NLS-SpCas9-(R221K N394K H840A)-GSGS-MLV RT pentamutant GS coon-opt SV40NLS-P2A-BSD	HEK3	+1 T to A #40	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGGCACGTGGCC (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGGCATCTCGTCGTCATG (SEQ ID NO: 5248)			GTCAACCAGTATCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A														
53A	HeLa	Chen2021 1323	dpSV 40NL S--Sp Cas9 R221 K N394 K H840 A)--S GGSx --XT EEN16 -SGG Sx3- MMLV RT pentamutan t GS ocean --tnS V40N LS--P 2A--B SD	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)				CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGCTCA (SEQ ID NO: 5248)		GTCAACCAAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021 1324	dpSV 40NL S--SpCas9 (R221K N394K H840A)--S GGSx3--M MLV RT pentamutan t GS ocean --tnS V40N LS--P 2A--B SD	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5241)	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)				CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGCTCA (SEQ ID NO: 5248)		GTCAACCAAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021 1324	dpSV 40NL S--SpCas9 (R221K N394K H840A)--S GGSx2--bp SV40NLS- SGGSx2-- MMLV RT pentamutan t GS ocean-sp SV40N LS--S -P2A--B GGSx2 ID Z-MM LV RT pentamutan t GS ocean --tnS V40N LS--P 2A--B SD	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5241)	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)				CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGCTCA (SEQ ID NO: 5248)		GTCAACCAAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

5'A	HeLa	Chen2021 1325	bpSV 40NL S--Sp Cas9 R221 K N394 K H840 A)--S GSGS pentamutant GS 3-bpS V40N LS--S GSGS ID 3-MM LV RT pentamutant GS occon --bnS V40N LS'--P 2A--B SD	bpSV40NL S--SpCas9 (R221K N394K H840A)--S GSGx3--bp SV40NLS-- GSGSx3-- MMLV RT pentamutant GS 3-bpS V40N LS--G --P2A--SS GSGx ID 3-MM LV RT pentamutant GS occon --bnS V40N LS'--P 2A--B SD	HEK3	HEK3	+1 T to A +80	+1 T to A +80	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGCTCA GTCTG (SEQ ID NO: 5248)	R	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)
5'A	HeLa	Chen2021 1326	bpSV 40NL S--Sp Cas9 R221 K N394 K H840 A)--S GSGS pentamutant GS 3-bpS V40N LS--S GSGS ID 3-MM LV RT pentamutant GS occon --bnS V40N LS'--P 2A--B SD	bpSV40NL S--SpCas9 (R221K N394K H840A)--S GSGx3--bp SV40NLS-- GSGSx3-- MMLV RT pentamutant GS 3-bpS V40N LS--G --P2A--SS GSGx ID 3-MM LV RT pentamutant GS occon --bnS V40N LS'--P 2A--B SD	HEK3	HEK3	+1 T to A +80	+1 T to A +80	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGCTCA GTCTG (SEQ ID NO: 5248)	R	GTCAACCAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1327	PE2–P2A–BSD [bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant–bpSV40NLS*–P2A–BSD]	PE2–P2A–BSD [bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant–bpSV40NLS*–P2A–BSD]	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1328	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1329	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–cmycNLS–P2A–BSD	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–cmycNLS–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1330	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–bpSV40NLS–P2A–BSD	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–bpSV40NLS–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1331	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1332	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx6–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx6–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1333	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx8–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx8–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1334	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx10–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx10–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1335	pbSV40NLS-SpCas9 (R221K N394K H840A)-GGSGG-M-MLV RT pentamutant GS pcon-pbSV40NLS-P2A-BSD	FANCF			GGAATCCGCTTCTGCAAGTTAGCACCC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGGGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)			GGGGGTCCCAGTGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1335	pbSV40NLS-SpCas9 (R221K N394K H840A)-GGSG4-5-GGSG2-M-MLV RT pentamutant GS pcon-pbSV40NLS-P2A-BSD	FANCF			GGAATCCGCTTCTGCAAGTTAGCACCC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGGGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)			GGGGGTCCCAGTGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

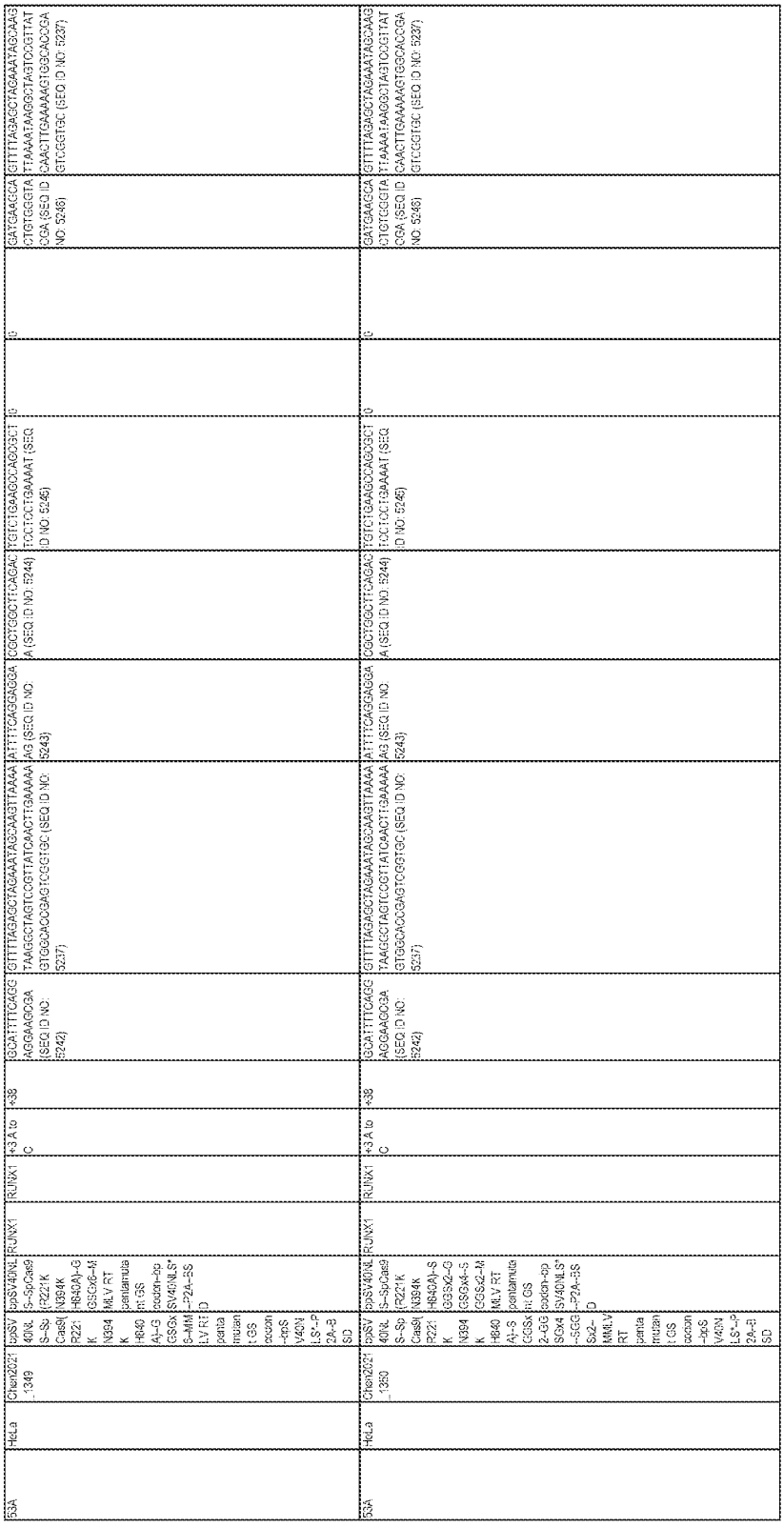
53A	HeLa	Chen2021_1337	bpSV40NLS--SpCas9(R221K N394K H840A)--SGGSx3--XTEN16--SGGSx3--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	bpSV40NLS--SpCas9(R221K N394K H840A)--SGGSx3--XTEN16--SGGSx3--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1338	bpSV40NLS--SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	bpSV40NLS--SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1339	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx3–bpSV40NLS–SGGSx3–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9 (R221K N394K H840A)–SGGSx3–bpSV40NLS–SGGSx3–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1340	bpSV40NLS–SpCas9(R221K N394K H840A)–GGSGx3–bpSV40NLS–GGSGx3–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9 (R221K N394K H840A)–GGSGx3–bpSV40NLS–GGSGx3–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1341	PE2--P2A--BSD [bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant--bpSV40NLS*--P2A--BSD]	PE2--P2A--BSD [bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant--bpSV40NLS*--P2A--BSD]	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1342	bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1343	bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--cmycNLS--P2A--BSD	bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--cmycNLS--P2A--BSD	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021:1344	bpSV40NLS-SpCas9-(H840A)-SGGSx2-KTEN16-SG35x2-M-MLV RT-panta mutant-1-3S-2-XTEN16-SGGSx2-MMLV-bpSV40NLS*-bpSV40NLS-P2A-BSD	RUNX1	F-3 A to C	+36	GCATTTTCAGGAGGAAGGAA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	10		GATCAACCGA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021:1345	bpSV40NLS-SpCas9-(R221K-N394K-H840A)-SGGSx2-KTEN16-SG35x2-M-MLV RT-panta mutant-1-3S-2-XTEN16-SGGSx2-MMLV-bpSV40NLS*-bpSV40NLS-P2A-BSD	RUNX1	F-3 A to C	+36	GCATTTTCAGGAGGAAGGAA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAT (SEQ ID NO: 5245)	10		GATCAACCGA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1346	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx6–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx6–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1347	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx8–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx8–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1348	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx10–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(R221K N394K H840A)–SGGSx10–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)


5'A
HeLa
Chen2021_1349
dpSV40NLS--SpCas9
dpSV40NLS--SpCas9 (R221K N394K H840A)-GGSGS--M-MLV RT pentamutant GS pcon --pvs V40N LS*--P2A--BD
RUNX1
+3 A to C
+38
GCATTTTCAGGAGGAAGGAGAAGGCA (SEQ ID NO: 5242)
GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
ATTTTCAGGAGGAAG (SEQ ID NO: 5243)
CGGCTGGCTTCAGAC (SEQ ID NO: 5244)
TGTCTGAAGCCAGGAGCTTCCTCCTGTGAAAAT (SEQ ID NO: 5245)
D
B
GATCGAACGCACTGTGGGTAACGA (SEQ ID NO: 5246)
GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

5'A	HeLa	Chen2021 1351	spSV40NLS-SpCas9 (T221K N394K H840A)-SGGSGGS-3-XTEN16-SV40NLS-SGGSGGS-PZA-SGSGSD MMLV RT pantamutant GS occon –tnS V40N LS–P2A–BSD	RUNX1	RUNX1	+3 A to C	+36	GCATTTTCAGGGAGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCTTCTCTGCTGAAAAAT (SEQ ID NO: 5245)			GATGAACGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
5'A	HeLa	Chen2021 1352	spSV40NLS-SpCas9 (T221K N394K H840A)-SGGSGGS-2-xtenSV40NLS-SGGSGGS-PZA-SGSGSD MMLV RT pantamutant GS occon –tnS V40N LS–P2A–BSD	RUNX1	RUNX1	+3 A to C	+36	GCATTTTCAGGGAGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCTTCTCTGCTGAAAAAT (SEQ ID NO: 5245)			GATGAACGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

5'A	HeLa	Chen2021 1353	bpSV40NL S–SpCas9 (T221K N394K H840A)-S GSGSx3–bpSV40NLS-SGGSGx3–MMLV RT (A1–S GSGSx3–codon-optV40N LS–SGSGx3–P2A–SD	RUNX1	RUNX1	+3 A to C	+36	GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCTTCTCTGCTGAAAAAT (SEQ ID NO: 5245)	10	8	GATGAACGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
5'A	HeLa	Chen2021 1354	bpSV40NL S–SpCas9 (T221K N394K H840A)-S GSGSx3–bpSV40NLS-SGGSGx3–MMLV RT (A1–S GSGSx3–codon-optV40N LS–SGSGSx3–P2A–SD	RUNX1	RUNX1	+3 A to C	+36	GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCTTCTCTGCTGAAAAAT (SEQ ID NO: 5245)	10	8	GATGAACGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1355	PE2–P2A–BSD [bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant–bpSV40NLS*–P2A–BSD]	PE2–P2A–BSD [bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant–bpSV40NLS*–P2A–BSD]	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1356	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1357	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–cmycNLS–P2A–BSD	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–cmycNLS–P2A–BSD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

G3A	HeLa	Chen2021 :359	bpSV 40NL S--Sp Cas9 (-H840 A)-S GGSG -SV40 NLS*- bpSV 40N LS*-b pSV4 0NLS -P2A -BSD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5286)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5286)	TGAGGGATCCAGATT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTC CTCATCTGGCCTGCAGA (SEQ ID NO: 5445)			GATGTACAG AGAGCCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
G3A	HeLa	Chen2021 :359	bpSV 40NL S--Sp Cas9 (-H840 A)-S GGSG -SV40 NLS*- bpSV 40N LS*-b pSV4 0NLS -P2A -BSD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5286)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5286)	TGAGGGATCCAGATT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTC CTCATCTGGCCTGCAGA (SEQ ID NO: 5445)			GATGTACAG AGAGCCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53A	HeLa	Chen2021_1383	pbSV40NLS-SpCas9 (T221K N394K R221K N394K R221K N394K K848A K848A A1-S G98Sx6–M MLV RT pantamutant GS codon-optSV40NLS*8–MMLV RT pantamutant GS codon-optSV40NLS*2A–BSD	VEGFA	VEGFA	+7 C to A	-57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCAGATCTATGCGCTT (SEQ ID NO: 5445)			GATGTACAG AGAGGCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1381	pbSV40NLS-SpCas9 (T221K N394K R221K N394K K848A K848A A1-S G98Sx6–MMLV RT pantamutant GS codon-optSV40NLS*2A–BSD	VEGFA	VEGFA	+7 C to A	-57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCAGATCTATGCGCTT (SEQ ID NO: 5445)			GATGTACAG AGAGGCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53A	HeLa	Chen2021_1382	pbSV40NLS-SpCas9 (R221K N394K K848A K848A GGSx10-MMLV RT pantamutant GS codon-optSV40NLS*10–MMLV RT pantamutant GS codon-optSV40NLS*2A–BSD	VEGFA	VEGFA	+7 C to A	-57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCAGATCTATGCGCTT (SEQ ID NO: 5445)			GATGTACAG AGAGGCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

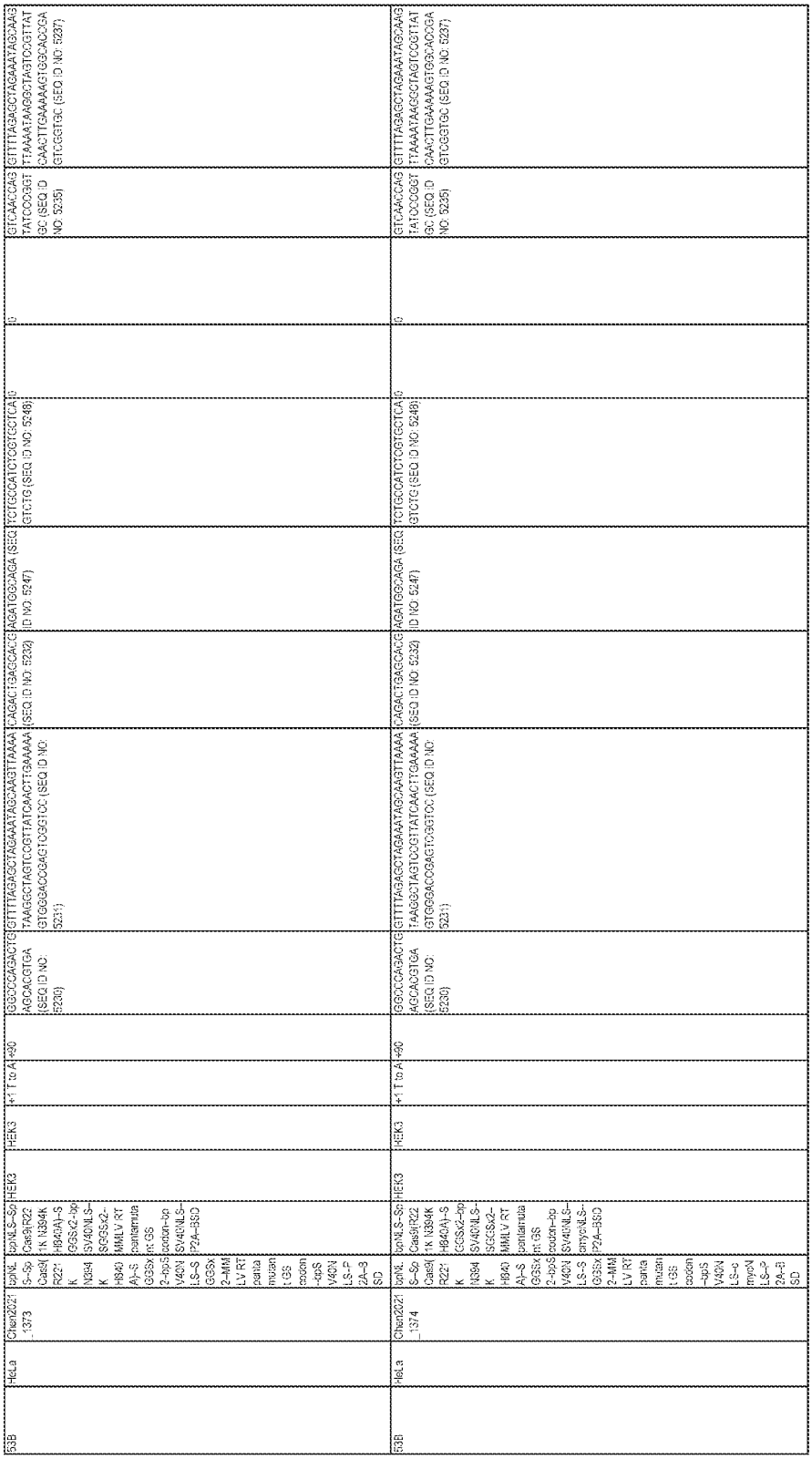
53A	Hela	Chen2021 1313	dpSV 40NL S--SpCas9 Cas9 R221 K N394 K H840 A1--S GGSGS--M LV RT pantaa mutan t GS ccon --apoS V40N LS*--P 2A--B SD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAG CCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA GTGGCAGCGAAGTCGGTTGC (SEQ ID NO: 5237)	TCTCCAGGCCAGA (SEQ ID NO: 5266)	TGACGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCG CTCATCTTGGCCTCGAGA (SEQ ID NO: 5445)			GATGTACAG AGAGCCCAG TGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTTGC (SEQ ID NO: 5237)
53A	Hela	Chen2021 1314	dpSV 40NL S--SpCas9 Cas9 R221K N394K H840A1--S GGSGS--M LV RT pantaa mutan t GS ccon --apoS V40N LS*--P 2A--B SD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAG CCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGGCCAGA (SEQ ID NO: 5266)	TGACGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCG CTCATCTTGGCCTCGAGA (SEQ ID NO: 5445)			GATGTACAG AGAGCCCAG TGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCCGTTGC (SEQ ID NO: 5237)

5'A	HDLB	Chen2021	bpSV40NLS-SpCas9-SpCas9 (T221K N394K H840A)-SGGSGGS-(G33S)2-K TEN16-SGGSGGS-M MLV RT pentamutant GS accon-tnS V40NLS*-P2A-GSGS-D MMLV RT pentamutant GS accon-tnS V40NLS*-P2A-BSD	VEGFA	147 C to A	457	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGAGCCAGA (SEQ ID NO: 5266)	TGACGGATCCACGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTCTCATCTTGGCCTGGCAGA (SEQ ID NO: 5445)		GATGTACAGAGAGGCCAGATGG (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
5'A	HDLB	Chen2021	bpSV40NLS-SpCas9-SpCas9 (T221K N394K H840A)-SGGSGGS-(G33S)2-bpSV40NLS-SGGSGGS-MMLV RT pentamutant GS accon-tnS V40NLS*-P2A-GSGS-D	VEGFA	147 C to A	457	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGAGCCAGA (SEQ ID NO: 5266)	TGACGGATCCACGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTCTCATCTTGGCCTGGCAGA (SEQ ID NO: 5445)		GATGTACAGAGAGGCCAGATGG (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

5'A	HeLa	Chen2021_1367	spSV40NLS-SpCas9-(T221K N394K R221K N394K H840A)-SGGSGG-spSV40NLS-SGGSGG-pentamutant GS 3-codon-opt SV40NLS-FP2A-SD 3-MMLV RT pentamutant GS 3'A	spSV40NLS-SpCas9-(T221K N394K R221K N394K H840A)-SGGSGG-spSV40NLS-SGGSGG-pentamutant GS 3-codon-opt SV40NLS-FP2A-SD 3-MMLV RT pentamutant GS 3'A	VEGFA	VEGFA	+7 C to A	+457	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGAGCCAGA (SEQ ID NO: 5266)	TGACGGATCCACGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTCTCATCTGGCCTGGCAGA (SEQ ID NO: 5445)		GATGTACAGAGACAGGCCCAGT GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
5'A	HeLa	Chen2021_1368	spSV40NLS-SpCas9-(T221K N394K R221K N394K H840A)-SGGSGG-spSV40NLS-SGGSGG-pentamutant GS 3-codon-opt SV40NLS-FP2A-SD 3-MMLV RT pentamutant GS 3'A	spSV40NLS-SpCas9-(T221K N394K R221K N394K H840A)-SGGSGG-spSV40NLS-SGGSGG-pentamutant GS 3-codon-opt SV40NLS-FP2A-SD 3-MMLV RT pentamutant GS 3'A	VEGFA	VEGFA	+7 C to A	+457	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGAGCCAGA (SEQ ID NO: 5266)	TGACGGATCCACGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTCTCATCTGGCCTGGCAGA (SEQ ID NO: 5445)		GATGTACAGAGACAGGCCCAGT GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1369	PE2--P2A--BSD [bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant--bpSV40NLS*--P2A--BSD]	PE2--P2A--BSD [bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant--bpSV40NLS*--P2A--BSD]	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1370	bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	bpSV40NLS--SpCas9(H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

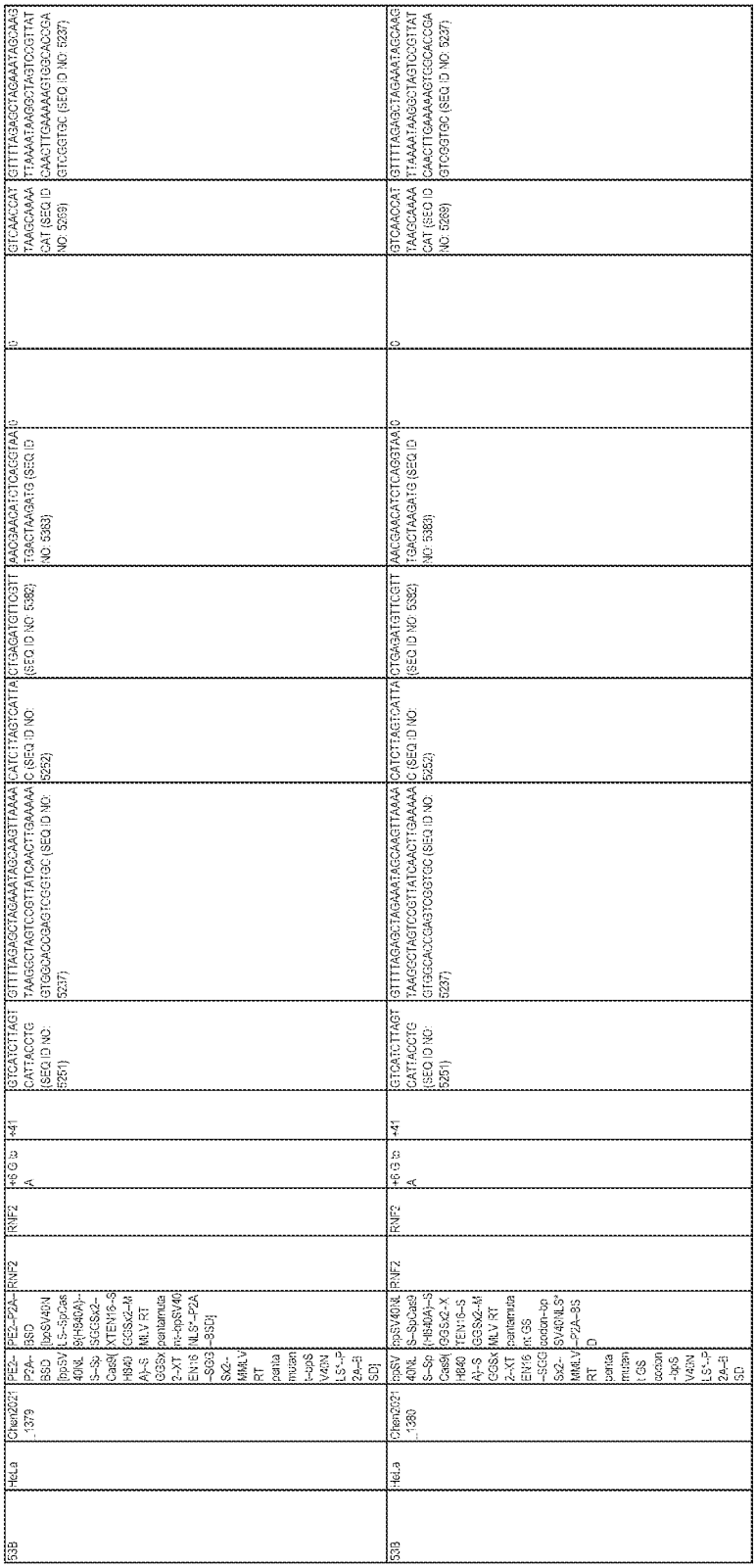
SSB	HeLa	Chen2021_1371	bpSV40NLS-SpCas9 (R221K N394K H840A)-SGSx2-XTEN16-SGSx2-MLV RT pentamutant GS 2-XT EN16 -SGS- Sx2-MMLV RT pentamutant GS 2-XT EN16 V40N LS--P 2A--B SD	HEK3	HEK3	+1 T to A +80	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGCTCA (SEQ ID NO: 5248)	R	GTCAACCAGTATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
SSB	HeLa	Chen2021_1372	bpSV40NLS-SpCas9 (R221K N394K H840A)-SGSx2-xv40NLS-SGSx2-MLV RT pentamutant GS 2-xv4 0NLS -SGS- Sx2-MMLV RT pentamutant GS 2-xv4 0NLS V40N LS--P 2A--B SD	HEK3	HEK3	+1 T to A +80	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGCTCA (SEQ ID NO: 5248)	R	GTCAACCAGTATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)



53B	HeLa	Chen2021_1373	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS–P2A–BSD	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS–P2A–BSD	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1374	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS–cmycNLS–P2A–BSD	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS–cmycNLS–P2A–BSD	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

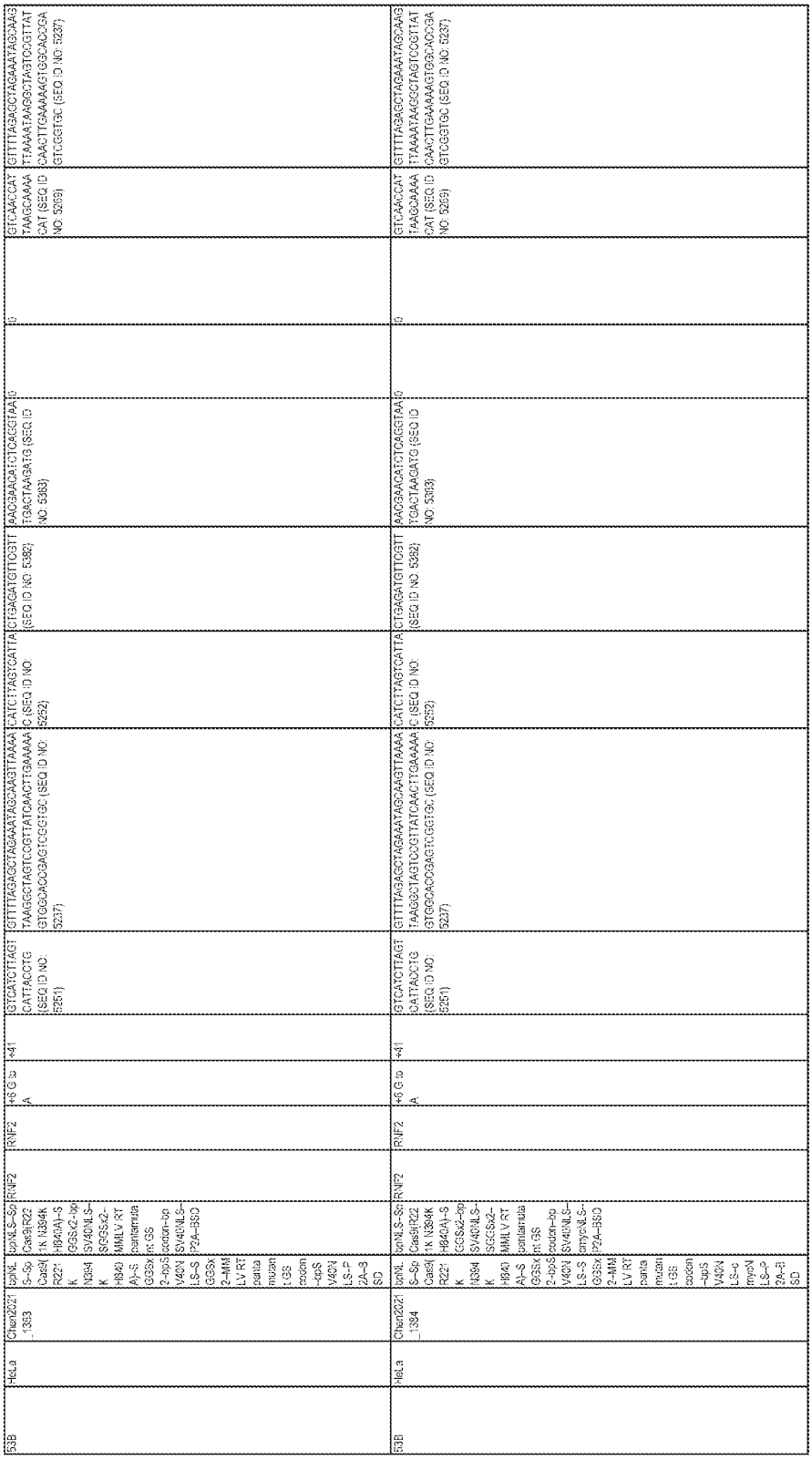
53B	HeLa	Chen2021_1375	bpNLS--SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS--bpSV40NLS--P2A--BSD	bpNLS--SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS--bpSV40NLS--P2A--BSD	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1376	SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS--P2A--BSD	SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS--P2A--BSD	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGGACCGAGTCGGTCC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5248)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1377	SpCas9 R61R2 221K N394K H840A)-SGGSx2-SV40NLS-SGGSx2-MMLV RT (D200N_L603W_T330P_T306K_W313F) LV RT panta mutant (GS codon-opt SV40NLS-nmycNLS-P2A-BSD	HEK3			GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGTCA (SEQ ID NO: 5248)			GTCAACCAG TATCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1378	SpCas9 R61R2 221K N394K H840A)-SGGSx2-SV40NLS-SGGSx2-MMLV RT (D200N_L603W_T330P_T306K_W313F) LV RT panta mutant (GS codon-opt SV40NLS-nmycNLS-P2A-BSD	HEK3			GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGA (SEQ ID NO: 5247)	TCTGCCATCTCGTCGTCA (SEQ ID NO: 5248)			GTCAACCAG TATCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)



53B		HeLa	Chen2021 -1379	PE2-P2A-BSD (pgSV40NL-SpCas9(H840A)-SGGSx2-XTEN16-SGGSx2-MMLV RT pantamutant-pgSV40NL-P2A-BSD)	PE2-P2A-BSD (pgSV40NL-SpCas9(H840A)-SGGSx2-XTEN16-SGGSx2-MMLV RT pantamutant-pgSV40NL-P2A-BSD)	RNF2	+6 G to A	+41	GTCATCTTAGT CATTACCTC (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGTGCACGCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAAG TGACTAAGATG (SEQ ID NO: 5383)	G		GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGCACGCGA GTCGGTGC (SEQ ID NO: 5237)
53B		HeLa	Chen2021 -1380	PE2-P2A-BSD (pgSV40NL-SpCas9(H840A)-SGGSx2-XTEN16-SGGSx2-MMLV RT pantamutant-pgSV40NL-P2A-BSD)	PE2-P2A-BSD (pgSV40NL-SpCas9(H840A)-SGGSx2-XTEN16-SGGSx2-MMLV RT pantamutant-pgSV40NL-P2A-BSD)	RNF2	+6 G to A	+41	GTCATCTTAGT CATTACCTC (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGTGCACGCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAAG TGACTAAGATG (SEQ ID NO: 5383)	G		GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGCGCACCGA GTCGGTGC (SEQ ID NO: 5237)

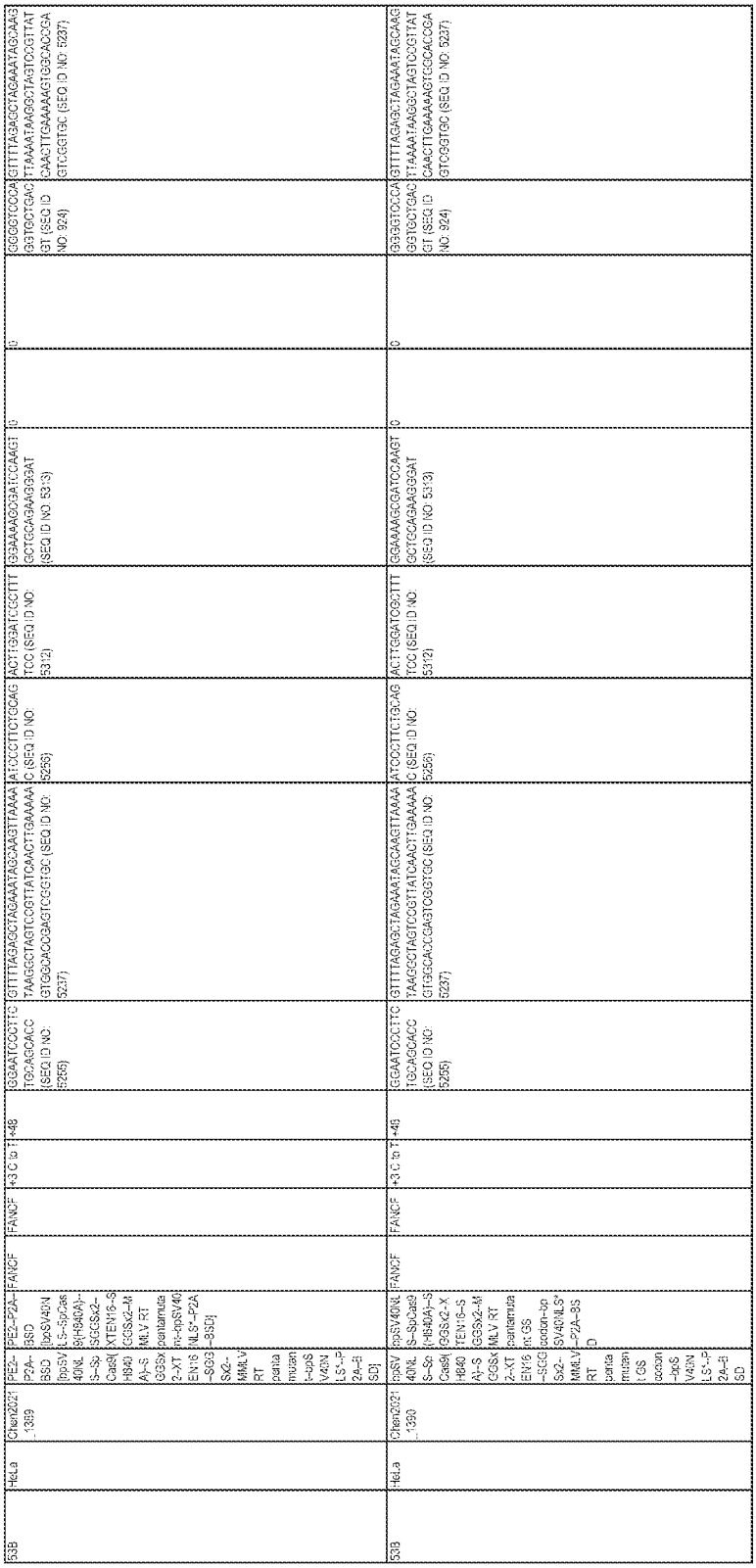
53B	HeLa	Chen2021_1381	bpSV40NLS--SpCas9(R221K N394K H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	bpSV40NLS--SpCas9(R221K N394K H840A)--SGGSx2--XTEN16--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1382	bpSV40NLS--SpCas9(R221K N394K H840A)--SGGSx2--sv40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	bpSV40NLS--SpCas9(R221K N394K H840A)--SGGSx2--sv40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS*--P2A--BSD	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)



53B	HeLa	Chen2021_1383	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–MMLV RT mutant GS codon-opt V403N LS–PZA–BSD	GTCATCTTAGTCATTACTCATTAGGT	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTACTCATTAGGT	CTTGAGATGTTGTGTT	AAGGAACATCTCAGGTAATG		GTCAACCATTAAGGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1384	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–MMLV RT mutant GS codon-opt V403N LS–PZA–BSD	GTCATCTTAGTCATTACTCATTAGGT	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTACTCATTAGGT	CTTGAGATGTTGTGTT	AAGGAACATCTCAGGTAATG		GTCAACCATTAAGGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1385	bpNLS--SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS--bpSV40NLS--P2A--BSD	bpNLS--SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS--bpSV40NLS--P2A--BSD	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1386	SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS--P2A--BSD	SpCas9(R221K N394K H840A)--SGGSx2--bpSV40NLS--SGGSx2--MMLV RT pentamutant GS codon--bpSV40NLS--P2A--BSD	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1387	SpCas9(R221K N394K H840A)-SGGSx2-bpSV40NLS-SGGSx2-MMLV RT pentamutant GS codon-bpSV40NLS-cmycNLS-P2A-BSD	SpCas9(R221K N394K H840A)-SGGSx2-bpSV40NLS-SGGSx2-MMLV RT pentamutant GS codon-bpSV40NLS-cmycNLS-P2A-BSD	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1388	SpCas9(R221K N394K H840A)-SGGSx2-bpSV40NLS-SGGSx2-MMLV RT pentamutant GS codon-bpSV40NLS-bpSV40NLS-P2A-BSD	SpCas9(R221K N394K H840A)-SGGSx2-bpSV40NLS-SGGSx2-MMLV RT pentamutant GS codon-bpSV40NLS-bpSV40NLS-P2A-BSD	RNF2	RNF2	+6 G to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	CTGAGATGTTCGTT (SEQ ID NO: 5382)	AACGAACATCTCAGGTAATGACTAAGATG (SEQ ID NO: 5383)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)



SSB	HeLa	Chen2021 -1389	PE2-P2A-BSD (pgRNA)	PE2-P2A-FANCF (nSV40NLS-SpCas9 (H840A)-SGGSx2-XTEN16-SGGSx2-MLV RT pentamutant-pgRNA)-SGGS-SGDJ	FANCF	+3 C to T 446	GGAATCCCTTC TGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC AC (SEQ ID NO: 5256)	ACTTGGATCCGTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5313)	10	0	GGGGGTCCCA GGTGGTGACT GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
SSB	HeLa	Chen2021 -1390	pgSV40NLS-SpCas9 (H840A)-SGGSx2-XTEN16-SGGSx2-MLV RT pentamutant-nt GS codon-opt-SV40NLS-P2A-BSD	FANCF	+3 C to T 446	GGAATCCCTTC TGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC AC (SEQ ID NO: 5256)	ACTTGGATCCGTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5313)	10	0	GGGGGTCCCA GGTGGTGACT GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1391	spSV40NLS-SpCas9-(T221K N394K H840A)-SGSx2-XTEN16-SGSx2-MLV RT pentamutant GS 2-XT EN16 SV40NLS*-SGS-P2A-SGSx2-MMLV RT pentamutant GS ocean-tnpB V40N LS*-P2A-SD	spSV40NLS-SpCas9-(T221K N394K H840A)-SGSx2-MLV RT pentamutant GS codon-opt SV40NLS*-SGS-P2A-SD	FANCF	FANCF	3'C to T	446	GGAATCCCTTCTGCAGCC	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGAATCCGTTTTCCC (SEQ ID NO: 5312)	GGAAAAGCGAATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	10	8	GGGGGTCCCAGGTGTGCTGAC (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1392	spSV40NLS-SpCas9-(T221K N394K H840A)-SGSx2-xtenNLS-SGSx2-MLV RT pentamutant GS codon-opt SV40NLS*-SGS-P2A-SD	spSV40NLS-SpCas9-(T221K N394K H840A)-SGSx2-xtenNLS-SGSx2-MLV RT pentamutant GS codon-opt SV40NLS*-SGS-P2A-SD	FANCF	FANCF	3'C to T	446	GGAATCCCTTCTGCAGCC	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGAATCCGTTTTCCC (SEQ ID NO: 5312)	GGAAAAGCGAATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	10	8	GGGGGTCCCAGGTGTGCTGAC (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1393	bpNLS–Sp Cas9(R22 Cas9(R22 K N394K R221K N394 K H840 A)–S G93A 2-4p5 V401I LS–S G93A 2–MM LV RT plant mutan t GS codon –apS V401I LS–P 2A–B SD	GGAATCCGTTC TGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG C (SEQ ID NO: 5256)	ACTTGAATCGCTTT TCC (SEQ ID NO: 5312)	GGAAAAGCGAATCCAAGT GCTGCAGAAGGGAT (SEQ ID NO: 5313)	10		GGGGGTCCCA GGTGCTGAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1394	bpNLS–Sp Cas9(R22 Cas9(R22 K N394K R221K N394 K H840 A)–S G93A 2-4p5 V401I LS–S G93A 2–MM LV RT plant mutan t GS codon –apS V401I LS–c myoN LS–S 2A–B SD	GGAATCCGTTC TGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAG C (SEQ ID NO: 5256)	ACTTGAATCGCTTT TCC (SEQ ID NO: 5312)	GGAAAAGCGAATCCAAGT GCTGCAGAAGGGAT (SEQ ID NO: 5313)	10		GGGGGTCCCA GGTGCTGAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

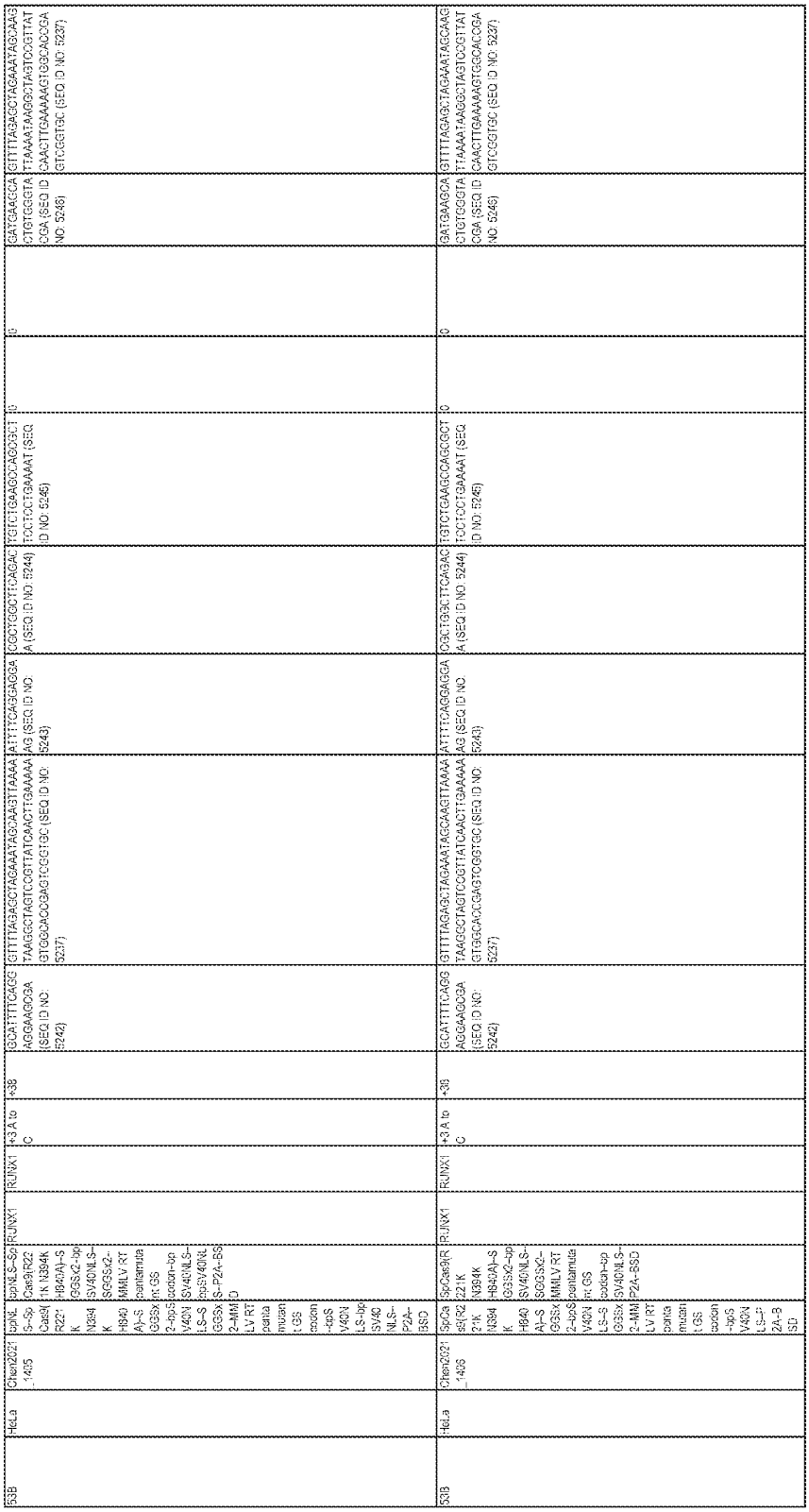
53B	HeLa	Chen2021_1395	spNLS–SpCas9(R221K N394K R22*-H840A)–SGGS(2-3p)–SV40NLS–MMLV RT pantamuta GSx codon-op V40N LS–SpSV40NLS–2-MMLV RT pantamutant GS codon –ipS V40N LS–bp SV40 NLS–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTC TGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGAATCCGTTTCC (SEQ ID NO: 5312)	GGAAAAGCGAATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGGTCCCAGGTGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1395	SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–MMLV RT pantamuta GSx codon-op V40N LS–SpSV40NLS–2-MMLV RT pantamutant GS codon –bpS V40N LS–P2A–BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTC TGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGAATCCGTTTCC (SEQ ID NO: 5312)	GGAAAAGCGAATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGGTCCCAGGTGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1397	SpCas9(R221K N394K H840A)-SGGSx2-bpSV40NLS-SGGSx2-MMLV RT pentamutant GS codon-bpSV40NLS-cmycNLS-P2A-BSD	SpCas9(R221K N394K H840A)-SGGSx2-bpSV40NLS-SGGSx2-MMLV RT pentamutant GS codon-bpSV40NLS-cmycNLS-P2A-BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1398	SpCas9(R221K N394K H840A)-SGGSx2-bpSV40NLS-SGGSx2-MMLV RT pentamutant GS codon-bpSV40NLS-bpSV40NLS-P2A-BSD	SpCas9(R221K N394K H840A)-SGGSx2-bpSV40NLS-SGGSx2-MMLV RT pentamutant GS codon-bpSV40NLS-bpSV40NLS-P2A-BSD	FANCF	FANCF	+3 C to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATCCCTTCTGCAGC (SEQ ID NO: 5256)	ACTTGGATCGCTTTTCC (SEQ ID NO: 5312)	GGAAAAGCGATCCAAGTGCTGCAGAAGGGAT (SEQ ID NO: 5313)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

S3B																
S3B	Hela	Chen2021:1399	PE2-P2A-BSD	PE2-P2A-BSD	RUNX1	RUNX1	+3 A to C	+36	GCATTTTCAGGAGGAAGGCAA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAAT (SEQ ID NO: 5245)	10	10	GATGAACCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
S3B	Hela	Chen2021:1400	rpSV40NLS–SpCas9 (rH840A)–SGGSx2–XTEN16–SGGSx2–MLV RT pantamutant RT with 5 additional mutations (rpSV40NLS–SpCas9 (rH840A)–SGGSx2–XTEN16–SGGSx2–MLV RT pantamutant RT with 5 additional mutations (rpSV40NLS–SpCas9 (rH840A)–SGGSx2–XTEN16–SGGSx2–MLV RT pantamutant RT with 5 additional mutations	RUNX1	RUNX1	+3 A to C	+36	GCATTTTCAGGAGGAAGGCAA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAAT (SEQ ID NO: 5245)	10	10	GATGAAGCA CTGTGGGTA CGA (SEQ ID NO: 5245)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

SSB	HeLa	Chen2021_1401	bpSV40NLS--SpCas9 (1221K N394K H840A)--SGSGx2--MMLV RT pentamutant GS 2--XT EN16 codon-opt SV40NLS*--SGSG--P2A--SD	RUNX1	RUNX1	+3 A to C	+36	GCATTTTCAGGGAGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCTTCTCTGCTGAAAAT (SEQ ID NO: 5245)			GATGAACCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
SSB	HeLa	Chen2021_1402	bpSV40NLS--SpCas9 (1221K N394K H840A)--SGSGx2--MMLV RT pentamutant GS 2--sv40 codon-opt SV40NLS*--SGSG--P2A--SD	RUNX1	RUNX1	+3 A to C	+36	GCATTTTCAGGGAGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGGTGGCTTCAGAC (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCTTCTCTGCTGAAAAT (SEQ ID NO: 5245)			GATGAACCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1403	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS–P2A–BSD	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS–P2A–BSD	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1404	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS–cmycNLS–P2A–BSD	bpNLS–SpCas9(R221K N394K H840A)–SGGSx2–bpSV40NLS–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS–cmycNLS–P2A–BSD	RUNX1	RUNX1	+3 A to C	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGCGCTTCCTCCTGAAAAT (SEQ ID NO: 5245)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)



53B	HeLa	Chen2021_1405	spNLS–SpCas9(R221K N394K H840A)–SGGS(x2–bpNLS–SV40NLS–P2A–BSD	RUNX1	RUNX1	+3 A to C	+3B	GCATTTTCAGGGAGGAAGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCTTCGCCTGCTTGAAAAT (SEQ ID NO: 5245)	TG		GATGAACCATCTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1405	SpCas9(R221K N394K H840A)–SGGS(x2–bpNLS–SV40NLS–P2A–BSD	RUNX1	RUNX1	+3 A to C	+3B	GCATTTTCAGGGAGGAAGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGGAGGAAG (SEQ ID NO: 5243)	CGCTGGCTTCAGACA (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCTTCGCCTGCTTGAAAAT (SEQ ID NO: 5245)	TG		GATGAACCATCTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1437	SpCas9 89/R2 221K N394K H840A SGGSx2 MMLV RT 2-aa pentamutation V4I/N L5~S codon-opt SV40NLS~ 2-AMM myoNLS~ LV RT P2A~BSD	SpCas9(R 221K N394K H840A)~S SGGSx2~bpNLS~ SV40NLS~ MMLV RT pentamutation rt GS codon-opt SV40NLS~ myoNLS~ P2A~BSD	RUNX1	RUNX1	+3 A to C	+3B	GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGAG (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAAT (SEQ ID NO: 5245)	10	10	GATGAAGCCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1439	SpCas9 89/R2 221K N394K H840A)~S SGGSx2~bpNLS~ SV40NLS~ MMLV RT 2-aa pentamutation V4I/N L5~S codon-opt SV40NLS~ 2-AMM bpSV40NLS~ LV RT P2A~BSD	SpCas9(R 221K N394K H840A)~S SGGSx2~bpNLS~ SV40NLS~ MMLV RT pentamutation rt GS codon-opt SV40NLS~ bpSV40NLS~ P2A~BSD	RUNX1	RUNX1	+3 A to C	+3B	GCATTTTCAGGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	AATTTTCAGGAGGAG (SEQ ID NO: 5243)	CGGCTGGCTTCAGAC A (SEQ ID NO: 5244)	TGTCTGAAGCCAGGGCT TCCTCCTGAAAAAT (SEQ ID NO: 5245)	10	10	GATGAAGCCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021:1419	PE2-P2A-BSD	PE2-P2A-DNMT1-BSD	DNMT1	1+6 to 1C	1+49	GATTGCTGGTGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAAACA (SEQ ID NO: 5262)	ACAGGGGTGAC (SEQ ID NO: 5422)	GTCAGGCCTGTTTCTGTGCACCAG (SEQ ID NO: 5423)	10	10	GCCCTTCAG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021:1419	PE2-P2A-BSD (bpSV40NLS–SpCas9 (1–1646)-SGGSx2-XTEN16–SGGSx2–MMLV RT pentamutant-bpSV40 NLS–P2A–BSD)	PE2-P2A-DNMT1-BSD (bpSV40NLS–SpCas9 (1–1646)-SGGSx2-XTEN16–SGGSx2–MMLV RT pentamutant-bpSV40 NLS–P2A–BSD)	DNMT1	1+6 to 1C	1+49	GATTGCTGGTGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAAACA (SEQ ID NO: 5262)	ACAGGGGTGAC (SEQ ID NO: 5422)	GTCAGGCCTGTTTCTGTGCACCAG (SEQ ID NO: 5423)	10	10	GCCCTTCAG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

SSB	HeLa	Chen2021 1411	dpSV40NLS--SpCas9 (1221K N394K H840A)--SGSG2--MLV RT pentamutant GS 2--NT EN16 codon-opt SV40NLS*--P2A--SD	DNMT1	DNMT1	1+6 G to C	149	GATTCGTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGTGTGCCAGGAA (SEQ ID NO: 5262)	ACACGCGTGAC (SEQ ID NO: 5422)	GTCACGCGTGTTTCTGG CACCAGG (SEQ ID NO: 5423)	10	10	GCCCTTCAG CTAAAATAAA TGG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
SSB	HeLa	Chen2021 1412	dpSV40NLS--SpCas9 (1221K N394K H840A)--SGSG2--MLV RT pentamutant GS 2--NT EN16 codon-opt SV40NLS*--P2A--SD	DNMT1	DNMT1	1+6 G to C	149	GATTCGTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGTGTGCCAGGAA (SEQ ID NO: 5262)	ACACGCGTGAC (SEQ ID NO: 5422)	GTCACGCGTGTTTCTGG CACCAGG (SEQ ID NO: 5423)	10	10	GCCCTTCAG CTAAAATAAA TGG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

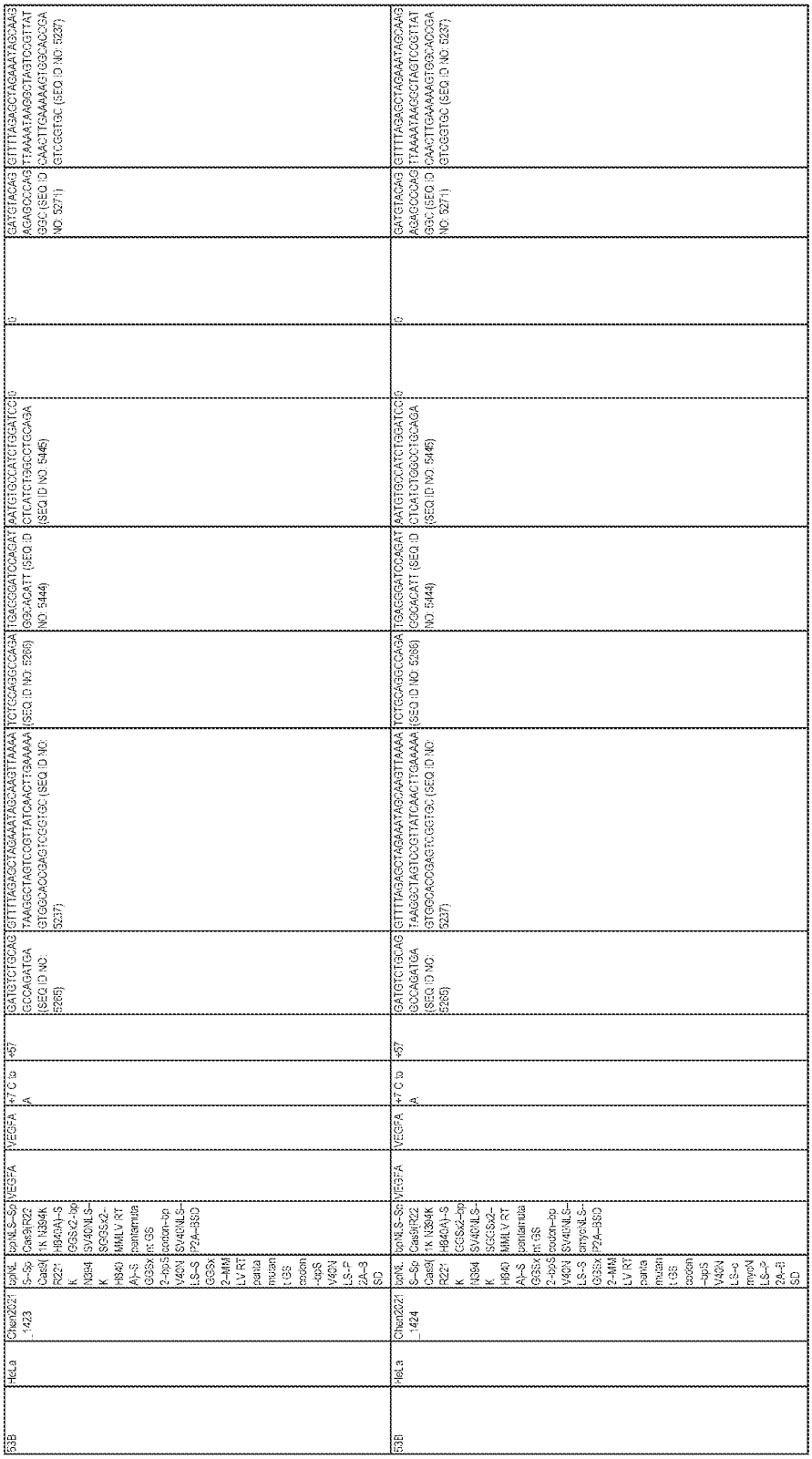
53B	HeLa	Chen2021_1413	bpNLS–Sp Cas9(R22 1N,N394K -H840A)–SGGSx2–bpNLS–SV40NLS–SGGSx2–MMLV RT (pantamuta mut GS codon-opt V40N LS–SGGSx2–MMLV RT (panta mutant GS codon-opt V40N LS–P2A–BSD	DNMT1	DNMT1	+6 G to C	+6 G to C	449	GATTCCTGTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCTGTGTGCCAGAAACAGAGGGCGTGAC (SEQ ID NO: 5262)	ACAGGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTTGGCACCAGG (SEQ ID NO: 5423)	G		GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1414	bpNLS–Sp Cas9(R22 1N,N394K -H840A)–SGGSx2–bpNLS–SV40NLS–SGGSx2–MMLV RT (pantamuta mut GS codon-opt V40N LS–SGGSx2–MMLV RT (panta mutant GS codon-opt V40N LS–P2A–BSD	DNMT1	DNMT1	+6 G to C	+6 G to C	449	GATTCCTGTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCTGTGTGCCAGAAACAGAGGGCGTGAC (SEQ ID NO: 5262)	ACAGGGCGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTTGGCACCAGG (SEQ ID NO: 5423)	G		GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1415	spNLS–SpCas9(R221K N394K H840A)–SGGS(x2–NLS–SV40 NLS–P2A–BSD	DNMT1	DNMT1	+6 G to C	449	GATTCGTGGTGCCAGAAA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTTGTGGCAGAA (SEQ ID NO: 5262)	ACACGGGTGAC (SEQ ID NO: 5422)	GTCACGCCTGTTTCTGG (SEQ ID NO: 5423)	G	G	GCCCTTCAG CTAAAATAAA TGG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1415	SpCas9(R221K N394K H840A)–SGGS(x2–NLS–SV40 NLS–P2A–BSD	DNMT1	DNMT1	+6 G to C	449	GATTCGTGGTGCCAGAAA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTGGTGCCAGAA (SEQ ID NO: 5262)	ACAGGGTGTGAC (SEQ ID NO: 5422)	GTCAGGCCTGTTTCTGG (SEQ ID NO: 5423)	G	G	GCCCTTCAG CTAAAATAAA TGG (SEQ ID NO: 5276)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

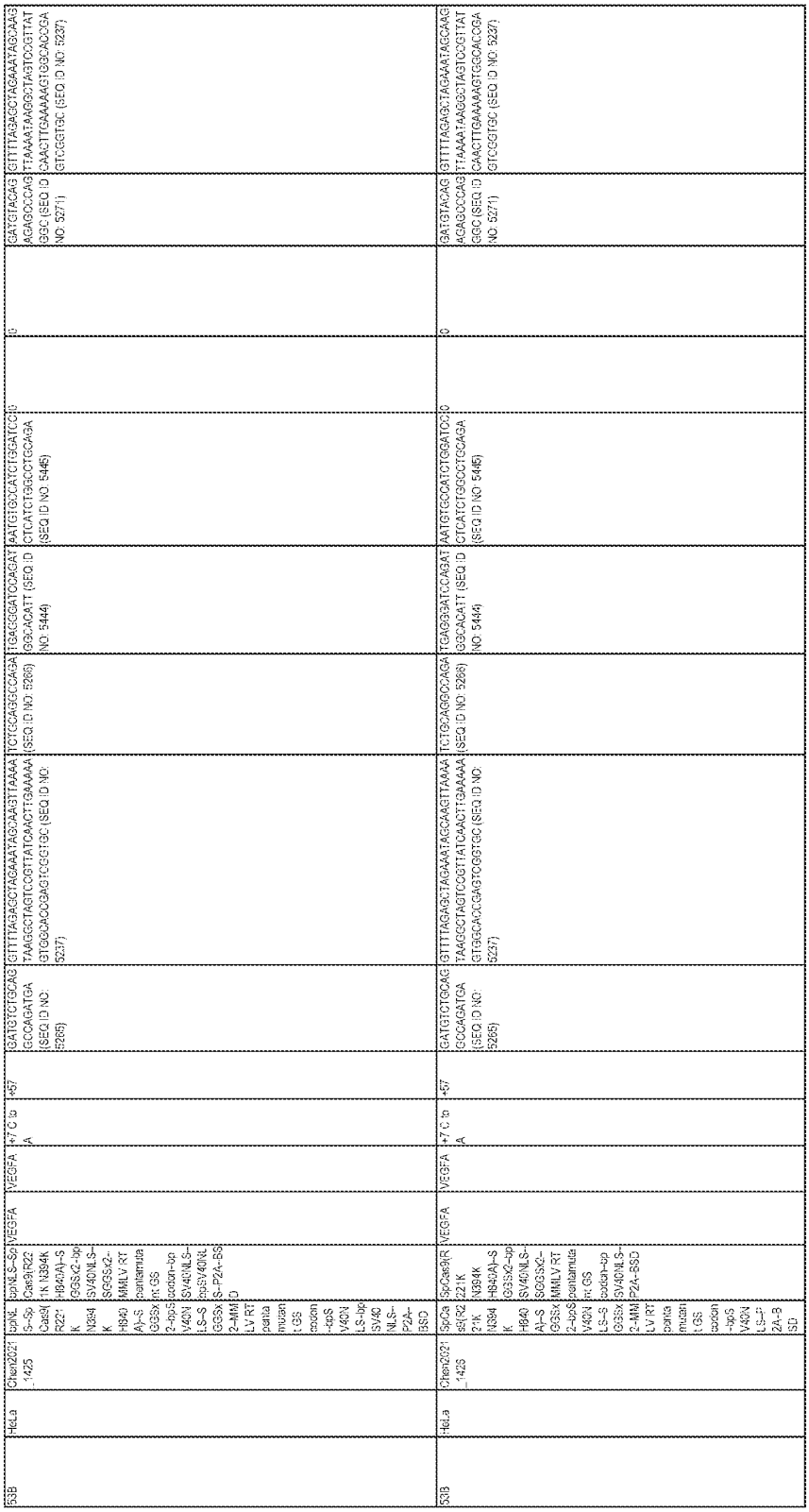
53B	HeLa	Chen2021_1417	SpCas9 89/R221K/N394K/H840A/S1335V/L1111R/D1135E/G1218R/E1219F/A1322P/R1333P/R1335Q/T1337R	SpCas9R221K/N394K/H840A/S1335V/L1111R/D1135E/G1218R/E1219F/A1322P/R1333P/R1335Q/T1337R	DNMT1	DNMT1	+6 G to C	+6 G to C	-449	-449	GATTCCTGTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTTGTGTGCCAGAAACAGGGCAGGTGAC (SEQ ID NO: 5262)	ACAGGGCAGGTGAC (SEQ ID NO: 5422)	GTCACCTGCCTGTTTCTTGGCACCAGG (SEQ ID NO: 5423)	10	10	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1418	SpCas9 89/R221K/N394K/H840A/S1335V/L1111R/D1135E/G1218R/E1219F/A1322P/R1333P/R1335Q/T1337R	SpCas9R221K/N394K/H840A/S1335V/L1111R/D1135E/G1218R/E1219F/A1322P/R1333P/R1335Q/T1337R	DNMT1	DNMT1	+6 G to C	+6 G to C	-449	-449	GATTCCTGTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CCTTGTGTGCCAGAAACAGGGCAGGTGAC (SEQ ID NO: 5262)	ACAGGGCAGGTGAC (SEQ ID NO: 5422)	GTCACCTGCCTGTTTCTTGGCACCAGG (SEQ ID NO: 5423)	10	10	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1419	PE2–P2A–BSD [bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant–bpSV40NLS*–P2A–BSD]	PE2–P2A–BSD [bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant–bpSV40NLS*–P2A–BSD]	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1420	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	bpSV40NLS–SpCas9(H840A)–SGGSx2–XTEN16–SGGSx2–MMLV RT pentamutant GS codon–bpSV40NLS*–P2A–BSD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCCTCATCTGGCCTGCAGA (SEQ ID NO: 5445)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

SSB	HeLa	Chen2021 1422	bpSV40NLS-SpCas9 (T221K N394K H840A)-SGGSx2-KTEN16-SGGSx2-MLV RT (A1-S285P)-SGGSx2-IDMMLV RT (D200N V428N L671P)-2A-BSD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGGCCAGA (SEQ ID NO: 5266)	TGACGGATCCACGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTCTCATCTTGGCCTGGCAGA (SEQ ID NO: 5445)		GATGTACAGAGAGGCCAGTGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
SSB	HeLa	Chen2021 1422	bpSV40NLS-SpCas9 (T221K N394K H840A)-SGGSx2-avidin-SV40NLS*-SGGSx2-MLV RT (A1-S285P)-SGGSx2-codon-opt (NLS SV40NLS*-SGGSx2-IDMMLV RT (D200N V428N L671P)-2A-BSD	VEGFA	VEGFA	+7 C to A	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTCCAGGCCAGA (SEQ ID NO: 5266)	TGACGGATCCACGATGGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTCTCATCTTGGCCTGGCAGA (SEQ ID NO: 5445)		GATGTACAGAGAGGCCAGTGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)



53B	HeLa	Chen2021_1423	bpNLS–Sp Cas9 (D10A-H840A-R221K-N394K-SGGSx2-bpNLS-SV40NLS-SGGSx2-MMLV RT (pantamuta mt GS 2-bp codon-opt V40N LS–SGGSx2–MMLV RT (pantamuta mutant GS codon-opt V40N LS–P2A–BSD	VEGFA	+7 C to A	-57	GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTC CTCATCTGGCCTGCAGA (SEQ ID NO: 5445)			GATGTACAG AGAGCCCAG TGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1424	bpNLS–Sp Cas9 (D10A-R221K-N394K-SGGSx2-bpNLS-SV40NLS-SGGSx2-MMLV RT (pantamuta mutant GS codon-opt V40N LS–smyoNLS–P2A–BSD	VEGFA	+7 C to A	-57	GATGTCTGCAG GCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGAT GGCACATT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTC CTCATCTGGCCTGCAGA (SEQ ID NO: 5445)			GATGTACAG AGAGCCCAG TGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)



53B	Hela	Chen2021_1425	spNLS–SpCas9(R221K N394K H840A)–SGGS(x2–bpNLS–SV40NLS–PZA–BSD	SpCas9(R221K N394K H840A)–SGGS(x2–bpNLS–SV40NLS–PZA–BSD	VEGFA	+7 C to A	457	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGCAGAT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTC (SEQ ID NO: 5445)		GATGTACAGAGCCAGATGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	Hela	Chen2021_1425	SpCas9(R221K N394K H840A)–SGGS(x2–bpNLS–SV40NLS–PZA–BSD	SpCas9(R221K N394K H840A)–SGGS(x2–bpNLS–SV40NLS–PZA–BSD	VEGFA	+7 C to A	457	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGGCCAGA (SEQ ID NO: 5266)	TGAGGGATCCAGATGCAGAT (SEQ ID NO: 5444)	AATGTGCCATCTGGATCCTC (SEQ ID NO: 5445)		GATGTACAGAGCCAGATGA (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53B	HeLa	Chen2021_1427	SpCas9 89/R221K/N394K/H840A/G1335A/R1335Q/T1337R	SpCas9(R221K/N394K/H840A)-SGGSx2-bpNMLV RT-GGSx2-contamundan RT GS codon-opV40N L5-SV40NLS-nmycNLS-P2A-BSD	VEGFA	HEK3	VEGFA	+7 C to A	+1 T to A	-57	-57	GATGTGTCTGCAGGCCAGATGA (SEQ ID NO: 5286)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGAGGCCAGA (SEQ ID NO: 5286)	TGAGGGATCCAGATGGCACAGAT (SEQ ID NO: 5444)	AATGTGGCATCTGGATCCAGTCTCATCTGGCCTGGCAGA (SEQ ID NO: 5445)	0	GATGTACAGAGACAGCCAGGTGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53B	HeLa	Chen2021_1428	SpCas9 89/R221K/N394K/H840A)-SGGSx2-bpNMLV RT-GGSx2-contamundan RT GS codon-opV40N L5-SV40NLS-nmycNLS-P2A-BSD	SpCas9 (R221K/N394K/H840A)-SGGSx2-bpNMLV RT-GGSx2-contamundan RT GS codon-opV40N L5-SV40NLS-nucSV40NLS-P2A-BSD	VEGFA	HEK3	VEGFA	+7 C to A	+1 T to A	-57	-57	GATGTGTCTGCAGGCCAGATGA (SEQ ID NO: 5286)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCTGCAGAGGCCAGA (SEQ ID NO: 5286)	TGAGGGATCCAGATGGCACAGAT (SEQ ID NO: 5444)	AATGTGGCATCTGGATCCAGTCTCATCTGGCCTGGCAGA (SEQ ID NO: 5445)	0	GATGTACAGAGACAGCCAGGTGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1429	PE2-P2A-BSD	PE2	HEK3	HEK3	+1 T to A			GGCCAGACTGAGCACGTGATGT	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTG (SEQ ID NO: 5232)	AGATGGCAGAGGAAG (SEQ ID NO: 5633)	TTCCTCTGCCATCTCGTGTG (SEQ ID NO: 5634)	0	0	0
52C	HeLa	Chen2021_1430	PEmax-P2A-BSD	PE2max	HEK3	HEK3	+1 T to A			GGCCAGACTGAGCACGTGATGT	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTG (SEQ ID NO: 5232)	AGATGGCAGAGGAAG (SEQ ID NO: 5633)	TTCCTCTGCCATCTCGTGTG (SEQ ID NO: 5634)	0	0	0
52C	HeLa	Chen2021_1431	PEmax-P2A-BSD	PE4max	HEK3	HEK3	+1 T to A			GGCCAGACTGAGCACGTGATGT	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTG (SEQ ID NO: 5232)	AGATGGCAGAGGAAG (SEQ ID NO: 5633)	TTCCTCTGCCATCTCGTGTG (SEQ ID NO: 5634)	0	0	0
52C	HeLa	Chen2021_1432	PEmax-P2A-BSD	PE4max-epeg	HEK3	HEK3	+1 T to A			GGCCAGACTGAGCACGTGATGT	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTG (SEQ ID NO: 5232)	AGATGGCAGAGGAAG (SEQ ID NO: 5633)	TTCCTCTGCCATCTCGTGTG (SEQ ID NO: 5634)	TCTCTCTGCCCCCCCCTGAACCCAGGCATAACCCCTCAAAGTCCGGGGGCAACCC (SEQ ID NO: 5606)	GGGTCAGGAGCGCCCCCCCCTGAACCCAGGCATAACCCCTCAAAGTCCGGGGGCAACCC (SEQ ID NO: 5606)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

52C	HeLa	Chen2021_1433	PE2-P2A-BSD	PE3	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAAAG (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTTG (SEQ ID NO: 5604)	0	0	GTCAACCAGTATCCCGTGT (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1434	PEmax-P2A-BSD	PE3max	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAAAG (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTTG (SEQ ID NO: 5604)	0	0	GTCAACCAGTATCCCGTGT (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1435	PEmax-P2A-BSD	PE5max	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAAAG (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTTG (SEQ ID NO: 5604)	0	0	GTCAACCAGTATCCCGTGT (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1436	PEmax-P2A-BSD-epeg	PE5max	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAAAG (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTTG (SEQ ID NO: 5604)	0	GGGTCAGGAGCCCCCCCCTGA TATCCCGGTACCCAGGGATAA CCCTCAAAGTCGGGGGGCAACC (SEQ ID NO: 5606)	0	GTCAACCAGTATCCCGTGT (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1437	PE2-P2A-BSD	PE2	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	0	0
52C	HeLa	Chen2021_1438	PEmax-P2A-BSD	PE2max	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	0	0
52C	HeLa	Chen2021_1439	PEmax-P2A-BSD	PE4max	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	0	0
52C	HeLa	Chen2021_1440	PEmax-P2A-BSD-epeg	PE4max	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	AACAAATCT (SEQ ID NO: 5607)	GGGTCAGGAGCCCCCCCCTGA ACCCAGGGATAA CCCTCAAAGTCGGGGGGCAACC (SEQ ID NO: 5606)	0	0
52C	HeLa	Chen2021_1441	PE2-P2A-BSD	PE3	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCGAT (SEQ ID NO: 5560)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1442	PEmax-P2A-BSD	PE3max	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCGAT (SEQ ID NO: 5560)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1443	PEmax-P2A-BSD	PE5max	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCGAT (SEQ ID NO: 5560)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1444	PEmax-P2A-BSD-epeg	PE5max	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	AACAAATCT (SEQ ID NO: 5607)	GGGTCAGGAGCCCCCCCCTGA ACCCAGGGATAA CCCTCAAAGTCGGGGGGCAACC (SEQ ID NO: 5606)	GACATCGATGTCCTCCCGAT (SEQ ID NO: 5560)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1445	PE2-P2A-BSD	PE2	RNF2	RNF2	+1 C to A		GTCATCTTAGTCATTGCTTACTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTACTG (SEQ ID NO: 5252)	ATGAGGTGTTCGTTTGAAC (SEQ ID NO: 5384)	AACGAACACCTCATGTAAA (SEQ ID NO: 5385)	0	0	0	0

52C	HeLa	Chen2021_1446	PEmax	PE2-P2A-BSD	RNF2	+1 C to A		GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5292)	ATGAGGTGTTCGTT TGACTAAGATG (SEQ ID NO: 5384)	AACGAACACCTCATGTAA TGACTAAGATG (SEQ ID NO: 5385)	0	0	0	0	0	0	0	0	0	0
52C	HeLa	Chen2021_1447	PEmax	PE4max-P2A-BSD	RNF2	+1 C to A		GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5292)	ATGAGGTGTTCGTT TGACTAAGATG (SEQ ID NO: 5384)	AACGAACACCTCATGTAA TGACTAAGATG (SEQ ID NO: 5385)	0	0	0	0	0	0	0	0	0	0
52C	HeLa	Chen2021_1448	PEmax	PE4max-P2A-ngpeg-BSD	RNF2	+1 C to A		GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5292)	ATGAGGTGTTCGTT TGACTAAGATG (SEQ ID NO: 5384)	AACGAACACCTCATGTAA TGACTAAGATG (SEQ ID NO: 5385)	TCATCTCT (SEQ ID NO: 5608)	TTGACGCGGTT CTATCTAGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	0	0	0	0	0	0	0	0	0
52C	HeLa	Chen2021_1449	PE2-P2A-BSD	PE3	RNF2	+1 C to A	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5292)	ATGAGGTGTTCGTT TGACTAAGATG (SEQ ID NO: 5384)	AACGAACACCTCATGTAA TGACTAAGATG (SEQ ID NO: 5385)	0	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	0	0	0	0	0
52C	HeLa	Chen2021_1450	PEmax	PE3max	RNF2	+1 C to A	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5292)	ATGAGGTGTTCGTT TGACTAAGATG (SEQ ID NO: 5384)	AACGAACACCTCATGTAA TGACTAAGATG (SEQ ID NO: 5385)	0	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	0	0	0	0	0
52C	HeLa	Chen2021_1451	PEmax	PE3max-P2A-BSD	RNF2	+1 C to A	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5292)	ATGAGGTGTTCGTT TGACTAAGATG (SEQ ID NO: 5384)	AACGAACACCTCATGTAA TGACTAAGATG (SEQ ID NO: 5385)	0	0	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	0	0	0	0	0
52C	HeLa	Chen2021_1452	PEmax	PE3max-P2A-ngpeg-BSD	RNF2	+1 C to A	-41	GTCATCTTAGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA C (SEQ ID NO: 5292)	ATGAGGTGTTCGTT TGACTAAGATG (SEQ ID NO: 5384)	AACGAACACCTCATGTAA TGACTAAGATG (SEQ ID NO: 5385)	TCATCTCT (SEQ ID NO: 5608)	TTGACGCGGTT CTATCTAGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	0	0	0	0	0	0
52C	HeLa	Chen2021_1453	PE2-P2A-BSD	PE2	FANCF	+5 G to T		TGGAATCCCTTC TGGAGCAACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT GCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	0	0	0	0	0	0	0	0
52C	HeLa	Chen2021_1454	PEmax	PE2max-P2A-BSD	FANCF	+5 G to T		TGGAATCCCTTC TGGAGCAACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT GCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	0	0	0	0	0	0	0	0
52C	HeLa	Chen2021_1455	PEmax	PE4max-P2A-BSD	FANCF	+5 G to T		TGGAATCCCTTC TGGAGCAACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT GCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	0	0	0	0	0	0	0	0
52C	HeLa	Chen2021_1456	PEmax	PE4max-ngpeg-BSD	FANCF	+5 G to T		TGGAATCCCTTC TGGAGCAACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT GCTGCAGAAGGGA (SEQ ID NO: 5612)	CAATCACT (SEQ ID NO: 5613)	TTGACGCGGTT CTATCTAGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	0	0	0	0	0	0	0	0	0
52C	HeLa	Chen2021_1457	PE2-P2A-BSD	PE3	FANCF	+5 G to T	-48	TGGAATCCCTTC TGCAAGCAACC (SEQ ID NO: 5256)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT GCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	0	GGGGGTCCCA GGTGCTGAC GT (SEQ ID NO: 924)	0	0	0	0	0
52C	HeLa	Chen2021_1458	PE2-P2A-BSD	PE3max	FANCF	+5 G to T	-48	TGGAATCCCTTC TCGCAGCAACC (SEQ ID NO: 5256)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT GCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	0	GGGGGTCCCA GGTGCTGAC GT (SEQ ID NO: 924)	0	0	0	0	0
52C	HeLa	Chen2021_1459	PEmax	PE5max-P2A-BSD	FANCF	+5 G to T	-48	TGGAATCCCTTC TGCAAGCAACC (SEQ ID NO: 5256)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT GCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	0	GGGGGTCCCA GGTGCTGAC GT (SEQ ID NO: 924)	0	0	0	0	0

52C	HeLa	Chen2021_1469	PEmax-P2A-BSD	PE5max-spcas	FANCF	RUNX1	+5 G to T	+48	GGAATCCGCTTC TGCACGCACC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC ATTTTCAGGAGGA (SEQ ID NO: 5910)	ACCTTGATCGCTTT TCC (SEQ ID NO: 5911)	GSAAAAGGGATCAAGGT GCTGCAGAAGGGA (SEQ ID NO: 5912)	CAATCACT (SEQ ID NO: 5913)	TTGACGCGGTT CTATCTATGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5608)	GGGGGTGCCA GGTGTGCTGAC GT (SEQ ID NO: 924)	0	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1481	PE2-P2A-BSD	PE2	RUNX1	RUNX1	+5 G to T	-38	CGATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCGTGAAAAT (SEQ ID NO: 5260)	CAATCACT (SEQ ID NO: 5913)	0	0	0	0
52C	HeLa	Chen2021_1482	PEmax-P2A-BSD	PE2max	RUNX1	RUNX1	+5 G to T	-38	CGATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCGTGAAAAT (SEQ ID NO: 5260)	CAATCACT (SEQ ID NO: 5913)	0	0	0	0
52C	HeLa	Chen2021_1483	PEmax-P2A-BSD	PE4max	RUNX1	RUNX1	+5 G to T	-38	CGATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCGTGAAAAT (SEQ ID NO: 5260)	CAATCACT (SEQ ID NO: 5913)	0	0	0	0
52C	HeLa	Chen2021_1484	PEmax-P2A-BSD	PE4max-spcas	RUNX1	RUNX1	+5 G to T	-38	CGATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCGTGAAAAT (SEQ ID NO: 5260)	AACTCTCT (SEQ ID NO: 5914)	TTGACGCGGTT CTATCTATGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	0	GATGAAGCGA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1485	PE2-P2A-BSD	PE3	RUNX1	RUNX1	+5 G to T	-38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCGTGAAAAT (SEQ ID NO: 5260)	AACTCTCT (SEQ ID NO: 5914)	0	0	GATGAAGCGA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1486	PEmax-P2A-BSD	PE3max	RUNX1	RUNX1	+5 G to T	-38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCGTGAAAAT (SEQ ID NO: 5260)	0	0	0	GATGAAGCGA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1487	PEmax-P2A-BSD	PE5max	RUNX1	RUNX1	+5 G to T	-38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCGTGAAAAT (SEQ ID NO: 5260)	0	0	0	GATGAAGCGA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1488	PEmax-P2A-BSD	PE5max-spcas	RUNX1	RUNX1	+5 G to T	-38	GCATTTTCAGG AGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGA AG (SEQ ID NO: 5243)	CGATTGCTTCAGAC A (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTT CCTCGTGAAAAT (SEQ ID NO: 5260)	AACTCTCT (SEQ ID NO: 5914)	TTGACGCGGTT CTATCTATGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	0	GATGAAGCGA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1468	PE2-P2A-BSD	PE2	DNMT1	DNMT1	+5 G to T	-38	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGTGCCAGA A (SEQ ID NO: 5915)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCTG ACCAGGA (SEQ ID NO: 5616)	0	0	0	0	0
52C	HeLa	Chen2021_1470	PEmax-P2A-BSD	PE2max	DNMT1	DNMT1	+5 G to T	-38	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGTGCCAGA A (SEQ ID NO: 5915)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCTG ACCAGGA (SEQ ID NO: 5616)	0	0	0	0	0
52C	HeLa	Chen2021_1471	PEmax-P2A-BSD	PE4max	DNMT1	DNMT1	+5 G to T	-38	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGTGCCAGA A (SEQ ID NO: 5915)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCTG ACCAGGA (SEQ ID NO: 5616)	0	0	0	0	0
52C	HeLa	Chen2021_1472	PEmax-P2A-BSD	PE4max-spcas	DNMT1	DNMT1	+5 G to T	-38	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGTGCCAGA A (SEQ ID NO: 5915)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCTG ACCAGGA (SEQ ID NO: 5616)	CCTCTTCT (SEQ ID NO: 5917)	TTGACGCGGTT CTATCTATGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	0	0	0
52C	HeLa	Chen2021_1473	PE2-P2A-BSD	PE3	DNMT1	DNMT1	+5 G to T	-49	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGTGCCAGA A (SEQ ID NO: 5915)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCTG ACCAGGA (SEQ ID NO: 5616)	0	GCCCTTCAG CTAAAATAAA (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

52C	HeLa	Chen2021_1474	PEmax-P2A-BSD	PE3max	DNMT1	DNMT1	+5 G to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1475	PEmax-P2A-BSD	PE5max	DNMT1	DNMT1	+5 G to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1476	PEmax-P2A-BSD	PE5max+epeg	DNMT1	DNMT1	+5 G to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	CCTCTTCT (SEQ ID NO: 5617)	TTGACGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAA (SEQ ID NO: 5609)	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1477	PE2-P2A-BSD	PE2	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	0	0
52C	HeLa	Chen2021_1478	PEmax-P2A-BSD	PE2max	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	0	0
52C	HeLa	Chen2021_1479	PEmax-P2A-BSD	PE4max	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	0	0
52C	HeLa	Chen2021_1480	PEmax-P2A-BSD	PE4max+epeg	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	ACAATCTC (SEQ ID NO: 5620)	TTGACGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAA (SEQ ID NO: 5609)	0	0
52C	HeLa	Chen2021_1481	PE2-P2A-BSD	PE3	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1482	PEmax-P2A-BSD	PE3max	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1483	PEmax-P2A-BSD	PE5max	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HeLa	Chen2021_1484	PEmax-P2A-BSD	PE5max+epeg	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	ACAATCTC (SEQ ID NO: 5620)	TTGACGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAA (SEQ ID NO: 5609)	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1485	PE2	PE2	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5604)	0	0	0	0
52C	HEK293T	Chen2021_1486	PEmax	PE2max	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5604)	0	0	0	0
52C	HEK293T	Chen2021_1487	PEmax	PE4max	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5604)	0	0	0	0

52C	HEK293T	Chen2021_1488	PEmax	PE4max +epeg	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5633)	TTCCTCTGGCCATCTCGTG (SEQ ID NO: 5634)	TCTCTCTC (SEQ ID NO: 5805)	GGGTCAGGAGCCCCCCCCTGA (SEQ ID NO: 5806)	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1489	PE2	PE3	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5633)	TTCCTCTGGCCATCTCGTG (SEQ ID NO: 5634)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1490	PEmax	PE3max	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5633)	TTCCTCTGGCCATCTCGTG (SEQ ID NO: 5634)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1491	PEmax	PE5max	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5633)	TTCCTCTGGCCATCTCGTG (SEQ ID NO: 5634)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1492	PEmax	PE5max +epeg	HEK3	+1 T to A	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5231)	CAGACTGAGCACGTGA (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5633)	TTCCTCTGGCCATCTCGTG (SEQ ID NO: 5634)	TCTCTCTC (SEQ ID NO: 5805)	GGGTCAGGAGCCCCCCCCTGA (SEQ ID NO: 5806)	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1493	PE2	PE2	EMX1	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1494	PE2max	PE2max	EMX1	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1495	PEmax	PE4max	EMX1	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1496	PEmax	PE4max +epeg	EMX1	+5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	AAGAATCT (SEQ ID NO: 5807)	GGGTCAGGAGCCCCCCCCTGA (SEQ ID NO: 5806)	0	0	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1497	PE2	PE3	EMX1	-5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	GGACATGGATGTCCTCCCCAT (SEQ ID NO: 5808)	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1498	PE3max	PE3max	EMX1	-5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	GGACATGGATGTCCTCCCCAT (SEQ ID NO: 5808)	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1499	PEmax	PE5max	EMX1	-5 G to T	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCATC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCT (SEQ ID NO: 5240)	0	0	GGACATGGATGTCCTCCCCAT (SEQ ID NO: 5808)	0	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

S2C	HEK293T	Chen2021_1503	PEmax	PE5max +epeg	EMX1	+5 G to T	+53	GAGTCCGAGCAGTAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAAGA (SEQ ID NO: 5238)	GAAGTGGTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGACGCACTTCTTCTTCTCCTCCGGGA (SEQ ID NO: 5240)	AACAATCT (SEQ ID NO: 5807)	GGGTGAGGAGCTCCCCCCCTGAAOCCAGGATAAOCTCTAAAGTCGCGAGGGGGCAACCC (SEQ ID NO: 5806)	GACATCGAT (GTCGTCCCCAT AT (SEQ ID NO: 5808)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
S2C	HEK293T	Chen2021_1501	PE2	PE2	RNF2	+1 C to A	+41	GTCATCTTATGTCATCCATCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCA (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5354)	AACGAACACCTCATGTAA (SEQ ID NO: 5365)	10	0	0	0
S2C	HEK293T	Chen2021_1502	PEmax	PE2max	RNF2	+1 C to A	+41	GTCATCTTATGTCATCCATCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCA (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5354)	AACGAACACCTCATGTAA (SEQ ID NO: 5365)	10	0	0	0
S2C	HEK293T	Chen2021_1503	PEmax	PE4max	RNF2	+1 C to A	+41	GTCATCTTATGTCATCCATCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCA (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5354)	AACGAACACCTCATGTAA (SEQ ID NO: 5365)	10	0	0	0
S2C	HEK293T	Chen2021_1504	PEmax	PE4max +epeg	RNF2	+1 C to A	+41	GTCATCTTATGTCATCCATCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCA (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5354)	AACGAACACCTCATGTAA (SEQ ID NO: 5365)	10	TTGAGGGCGGTT GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5809)	0	0
S2C	HEK293T	Chen2021_1505	PE2	PE3	RNF2	+1 C to A	+41	GTCATCTTATGTCATCCATCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCA (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5354)	AACGAACACCTCATGTAA (SEQ ID NO: 5365)	10	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
S2C	HEK293T	Chen2021_1506	PEmax	PE3max	RNF2	+1 C to A	+41	GTCATCTTATGTCATCCATCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCA (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5354)	AACGAACACCTCATGTAA (SEQ ID NO: 5365)	10	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
S2C	HEK293T	Chen2021_1507	PEmax	PE5max +epeg	RNF2	+1 C to A	+41	GTCATCTTATGTCATCCATCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCA (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5354)	AACGAACACCTCATGTAA (SEQ ID NO: 5365)	10	0	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
S2C	HEK293T	Chen2021_1508	PEmax	PE5max +epeg	RNF2	+1 C to A	+41	GTCATCTTATGTCATCCATCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAATCA (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5354)	AACGAACACCTCATGTAA (SEQ ID NO: 5365)	10	ITTGAGGGCGGTT GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5809)	GTCAACCAT TAAGCAAAA CAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
S2C	HEK293T	Chen2021_1509	PE2	FANCF	+5 G to T		GGAATCCGCTTCTGCAGCAGCC (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCCTTCTGCAGC (SEQ ID NO: 5610)	AGCTTGATCGCTTTACCTCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT (SEQ ID NO: 5612)	10	0	0	0
S2C	HEK293T	Chen2021_1510	PEmax	PE2max	FANCF	+5 G to T		GGAATCCGCTTCTGCAGCAGCC (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCCTTCTGCAGC (SEQ ID NO: 5610)	AGCTTGATCGCTTTACCTCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT (SEQ ID NO: 5612)	10	0	0	0
S2C	HEK293T	Chen2021_1511	PEmax	PE4max	FANCF	+5 G to T		GGAATCCGCTTCTGCAGCAGCC (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCCTTCTGCAGC (SEQ ID NO: 5610)	AGCTTGATCGCTTTACCTCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT (SEQ ID NO: 5612)	10	0	0	0
S2C	HEK293T	Chen2021_1512	PEmax	PE4max +epeg	FANCF	+5 G to T		GGAATCCGCTTCTGCAGCAGCC (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCCTTCTGCAGC (SEQ ID NO: 5610)	AGCTTGATCGCTTTACCTCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT (SEQ ID NO: 5612)	10	TTGAGGGCGGTT GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5809)	0	0
S2C	HEK293T	Chen2021_1513	PE2	PE3	FANCF	+5 G to T	+48	GGAATCCGCTTCTGCAGCAGCC (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCCTTCTGCAGC (SEQ ID NO: 5610)	AGCTTGATCGCTTTACCTCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT (SEQ ID NO: 5612)	10	0	GGGGGTCCCCA GGGTGGTGAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

52C	HEK293T	Chen2021_1514	PEmax	PE3max	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1515	PEmax	PE5max	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1516	PEmax	PE5max +epeg	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	CAATCACT (SEQ ID NO: 5613)	TTGACGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAA (SEQ ID NO: 5609)	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1517	PE2	PE2	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
52C	HEK293T	Chen2021_1518	PEmax	PE2max	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
52C	HEK293T	Chen2021_1519	PEmax	PE4max	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
52C	HEK293T	Chen2021_1520	PEmax	PE4max +epeg	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	AACTCTCT (SEQ ID NO: 5614)	TTGACGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAA (SEQ ID NO: 5609)	0	0
52C	HEK293T	Chen2021_1521	PE2	PE3	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1522	PEmax	PE3max	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1523	PEmax	PE5max	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1524	PEmax	PE5max +epeg	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	AACTCTCT (SEQ ID NO: 5614)	TTGACGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAAA (SEQ ID NO: 5609)	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1525	PE2	PE2	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	0	0
52C	HEK293T	Chen2021_1526	PEmax	PE2max	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	0	0
52C	HEK293T	Chen2021_1527	PEmax	PE4max	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	0	0

52C	HEK293T	Chen2021_1528	PEmax	PE4max +epeg	DNMT1	DNMT1	+5 G to T		GATTCGTGGTG CCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCGTGGTGCCAGA A (SEQ ID NO: 5615)	ACAGTGGTGAC ACCAGGA (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCG ACCAGGA (SEQ ID NO: 5616)	CCTCTTCT (SEQ ID NO: 5617)	TTGACGCGGTT CTATCTAGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	0	GCCCTTCAG CTAAAATAAA GG (SEQ ID NO: 5270)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1529	PE2	PE3	DNMT1	DNMT1	+5 G to T	+48	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGGTGCCAGA A (SEQ ID NO: 5615)	ACAGTGGTGAC ACCAGGA (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCG ACCAGGA (SEQ ID NO: 5616)	CCTCTTCT (SEQ ID NO: 5617)	0	0	GCCCTTCAG CTAAAATAAA GG (SEQ ID NO: 5270)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1530	PEmax	PE3max	DNMT1	DNMT1	+5 G to T	+48	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGGTGCCAGA A (SEQ ID NO: 5615)	ACAGTGGTGAC ACCAGGA (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCG ACCAGGA (SEQ ID NO: 5616)	CCTCTTCT (SEQ ID NO: 5617)	0	0	GCCCTTCAG CTAAAATAAA GG (SEQ ID NO: 5270)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1531	PEmax	PE5max	DNMT1	DNMT1	+5 G to T	+48	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGGTGCCAGA A (SEQ ID NO: 5615)	ACAGTGGTGAC ACCAGGA (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCG ACCAGGA (SEQ ID NO: 5616)	CCTCTTCT (SEQ ID NO: 5617)	0	0	GCCCTTCAG CTAAAATAAA GG (SEQ ID NO: 5270)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1532	PEmax	PE5max +epeg	DNMT1	DNMT1	+5 G to T	+48	GATTCCTGGTG CCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGTGGTGCCAGA A (SEQ ID NO: 5615)	ACAGTGGTGAC ACCAGGA (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCG ACCAGGA (SEQ ID NO: 5616)	CCTCTTCT (SEQ ID NO: 5617)	TTGACGCGGTT CTATCTAGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	0	GCCCTTCAG CTAAAATAAA GG (SEQ ID NO: 5270)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1533	PE2	PE2max	VEGFA	VEGFA	+5 G to T		CATCTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGAA (SEQ ID NO: 5618)	TCAGTGGCTCCAGAT AGGCACATT (SEQ ID NO: 5440)	AATGTGGCCATCTGGAGC TCATCATCTGGGCTGCAG (SEQ ID NO: 5619)	0	0	0	0	0	0
52C	HEK293T	Chen2021_1534	PEmax	PE2max	VEGFA	VEGFA	+5 G to T		GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGAA (SEQ ID NO: 5618)	TCAGTGGCTCCAGAT AGGCACATT (SEQ ID NO: 5440)	AATGTGGCCATCTGGAGC TCATCATCTGGGCTGCAG (SEQ ID NO: 5619)	0	0	0	0	0	0
52C	HEK293T	Chen2021_1535	PEmax	PE4max	VEGFA	VEGFA	+5 G to T		GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGAA (SEQ ID NO: 5618)	TCAGTGGCTCCAGAT AGGCACATT (SEQ ID NO: 5440)	AATGTGGCCATCTGGAGC TCATCATCTGGGCTGCAG (SEQ ID NO: 5619)	0	0	0	0	0	0
52C	HEK293T	Chen2021_1536	PEmax	PE4max +epeg	VEGFA	VEGFA	+5 G to T		GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGAA (SEQ ID NO: 5618)	TCAGTGGCTCCAGAT AGGCACATT (SEQ ID NO: 5440)	AATGTGGCCATCTGGAGC TCATCATCTGGGCTGCAG (SEQ ID NO: 5619)	ACAATCTC (SEQ ID NO: 5620)	TTGACGCGGTT CTATCTAGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	0	0	0	0	0
52C	HEK293T	Chen2021_1537	PE2	PE3	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGAA (SEQ ID NO: 5618)	TCAGTGGCTCCAGAT AGGCACATT (SEQ ID NO: 5440)	AATGTGGCCATCTGGAGC TCATCATCTGGGCTGCAG (SEQ ID NO: 5619)	0	0	GATGTACAG AGAGCGCAG GGC (SEQ ID NO: 5271)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1538	PEmax	PE3max	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGAA (SEQ ID NO: 5618)	TCAGTGGCTCCAGAT AGGCACATT (SEQ ID NO: 5440)	AATGTGGCCATCTGGAGC TCATCATCTGGGCTGCAG (SEQ ID NO: 5619)	0	0	GATGTACAG AGAGCGCAG GGC (SEQ ID NO: 5271)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1539	PEmax	PE5max	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGAA (SEQ ID NO: 5618)	TCAGTGGCTCCAGAT AGGCACATT (SEQ ID NO: 5440)	AATGTGGCCATCTGGAGC TCATCATCTGGGCTGCAG (SEQ ID NO: 5619)	0	0	GATGTACAG AGAGCGCAG GGC (SEQ ID NO: 5271)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52C	HEK293T	Chen2021_1540	PEmax	PE5max +epeg	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAG GCCACGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAAG TGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGAA (SEQ ID NO: 5618)	TCAGTGGCTCCAGAT AGGCACATT (SEQ ID NO: 5440)	AATGTGGCCATCTGGAGC TCATCATCTGGGCTGCAG (SEQ ID NO: 5619)	ACAATCTC (SEQ ID NO: 5620)	TTGACGCGGTT CTATCTAGTTAC GCGTTAAACCAA CTAGAAA (SEQ ID NO: 5609)	GATGTACAG AGAGCGCAG GGC (SEQ ID NO: 5271)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

53C	Hela	Chen2021_1541	PE2-P2A-BSD	HEK3	HEK3	+1 T to A	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAAAC (SEQ ID NO: 5232)	AGATGGCAGAGTGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGTG (SEQ ID NO: 5604)	0	0	0	0
53C	Hela	Chen2021_1542	PEmax-P2A-BSD	HEK3	HEK3	+1 T to A	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAAAC (SEQ ID NO: 5232)	AGATGGCAGAGTGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGTG (SEQ ID NO: 5604)	0	0	0	0
53C	Hela	Chen2021_1543	PE2-P2A-BSD	HEK3	HEK3	+1 T to A +40	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAAAC (SEQ ID NO: 5232)	AGATGGCAGAGTGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGTG (SEQ ID NO: 5604)	0	0	GTCAACCAAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1544	PEmax-P2A-BSD	HEK3	HEK3	+1 T to A +40	GGCCCAGACTG AGCACGTGA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACGTGAAAAC (SEQ ID NO: 5232)	AGATGGCAGAGTGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGTG (SEQ ID NO: 5604)	0	0	GTCAACCAAG TATCCCGGT GC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1545	PE2-P2A-BSD	EMX1	EMX1	+5 G to T	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCCCATCG (SEQ ID NO: 5240)	GTGATGGGAAGCACTTCT (SEQ ID NO: 5241)	0	0	0	0
53C	Hela	Chen2021_1546	PEmax-P2A-BSD	EMX1	EMX1	+5 G to T	GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCCCATCG (SEQ ID NO: 5240)	GTGATGGGAAGCACTTCT (SEQ ID NO: 5241)	0	0	0	0
53C	Hela	Chen2021_1547	PE2-P2A-BSD	EMX1	EMX1	+5 G to T	+53 GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCCCATCG (SEQ ID NO: 5240)	GTGATGGGAAGCACTTCT (SEQ ID NO: 5241)	0	0	GACATCGAT GTCCTCGCC AT (SEQ ID NO: 5558)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1548	PEmax-P2A-BSD	EMX1	EMX1	+5 G to T	+53 GAGTCCGAGCA GAAGAAGAA (SEQ ID NO: 5238)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5239)	GAAGTGGTCCCATCG (SEQ ID NO: 5240)	GTGATGGGAAGCACTTCT (SEQ ID NO: 5241)	0	0	GACATCGAT GTCCTCGCC AT (SEQ ID NO: 5558)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1549	PE2-P2A-BSD	RNF2	RNF2	+1 C to A	GTCATCTTATGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA ATGAGGTGTTCGTT (SEQ ID NO: 5252)	AACGAACACCTCATGTAA TGACTTAAGATG (SEQ ID NO: 5365)	0	0	0	0
53C	Hela	Chen2021_1550	PEmax-P2A-BSD	RNF2	RNF2	+1 C to A	GTCATCTTATGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA ATGAGGTGTTCGTT (SEQ ID NO: 5252)	AACGAACACCTCATGTAA TGACTTAAGATG (SEQ ID NO: 5365)	0	0	0	0
53C	Hela	Chen2021_1551	PE2-P2A-BSD	RNF2	RNF2	+1 C to A	+41 GTCATCTTATGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA ATGAGGTGTTCGTT (SEQ ID NO: 5252)	AACGAACACCTCATGTAA TGACTTAAGATG (SEQ ID NO: 5365)	0	GTCAAGCAT TAAGCAAAA CAT (SEQ ID NO: 5268)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1552	PEmax-P2A-BSD	RNF2	RNF2	+1 C to A	+41 GTCATCTTATGT CATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTA ATGAGGTGTTCGTT (SEQ ID NO: 5252)	AACGAACACCTCATGTAA TGACTTAAGATG (SEQ ID NO: 5365)	0	GTCAAGCAT TAAGCAAAA CAT (SEQ ID NO: 5268)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1553	PE2-P2A-BSD	FANCF	FANCF	+5 G to T	GGAATCCGCTTC TCGACGCAGCC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5227)	TCCGCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT (SEQ ID NO: 5612)	0	0	0	0
53C	Hela	Chen2021_1554	PEmax-P2A-BSD	FANCF	FANCF	+5 G to T	GGAATCCGCTTC TCGACGCAGCC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5227)	TCCGCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT (SEQ ID NO: 5612)	0	0	0	0
53C	Hela	Chen2021_1555	PE2-P2A-BSD	FANCF	FANCF	+5 G to T	+48 GGAATCCGCTTC TCGACGCAGCC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTT TCC (SEQ ID NO: 5611)	GGAAAAGCGGATCAAGGT (SEQ ID NO: 5612)	0	GGGGGTCCCA GCGTGCTGAC GT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

53C	Hela	Chen2021_1556	PEmax–P2A–BSD	PE5max	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1557	PE2–P2A–BSD	PE4	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
53C	Hela	Chen2021_1558	PEmax–P2A–BSD	PE4max	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0
53C	Hela	Chen2021_1559	PE2–P2A–BSD	PE5	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1560	PEmax–P2A–BSD	PE5max	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	GATGAAGCACTGTGGGTACGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1561	PE2–P2A–BSD	PE4	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	0	0
53C	Hela	Chen2021_1562	PEmax–P2A–BSD	PE4max	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	0	0
53C	Hela	Chen2021_1563	PE2–P2A–BSD	PE5	DNMT1	DNMT1	+5 G to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1564	PEmax–P2A–BSD	PE5max	DNMT1	DNMT1	+5 G to T	+49	GATTCCTGGTGCCAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGAA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	GTCACCACTGTTTCTGGCACCAGGA (SEQ ID NO: 5616)	0	0	GCCCTTCAGCTAAAATAAAGG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1565	PE2–P2A–BSD	PE4	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	0	0
53C	Hela	Chen2021_1566	PEmax–P2A–BSD	PE4max	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	0	0
53C	Hela	Chen2021_1567	PE2–P2A–BSD	PE5	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	Hela	Chen2021_1568	PEmax–P2A–BSD	PE5max	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCCAGATGA (SEQ ID NO: 5265)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGATGGCACATT (SEQ ID NO: 5440)	AATGTGCCATCTGGAGCACTCATCTGGCCTGCAG (SEQ ID NO: 5619)	0	0	GATGTACAGAGAGCCCAGGGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1569	PE2	PE4	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5604)	0	0	0	0
53C	HEK293T	Chen2021_1570	PEmax	PE4max	HEK3	HEK3	+1 T to A		GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5604)	0	0	0	0

53C	HEK293T	Chen2021_1571	PE2	PE5	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5604)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1572	PEmax	PE5max	HEK3	HEK3	+1 T to A	+90	GGCCCAGACTGAGCACGTGA (SEQ ID NO: 5230)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CAGACTGAGCACG (SEQ ID NO: 5232)	AGATGGCAGAGGAA (SEQ ID NO: 5603)	TTCCTCTGCCATCTCGTGCTCAGTCTG (SEQ ID NO: 5604)	0	0	GTCAACCAGTATCCCGGTGC (SEQ ID NO: 5235)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1573	PE2	PE4	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
53C	HEK293T	Chen2021_1574	PEmax	PE4max	EMX1	EMX1	+5 G to T		GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	0	0
53C	HEK293T	Chen2021_1575	PE2	PE5	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1576	PEmax	PE5max	EMX1	EMX1	+5 G to T	+53	GAGTCCGAGCAGAAGAAGAA (SEQ ID NO: 5236)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCGAGCAGAAGAA (SEQ ID NO: 5238)	GAAGTGCTCCCATCAC (SEQ ID NO: 5239)	GTGATGGGAGCACTTCTTCTTCTGCTCGGA (SEQ ID NO: 5240)	0	0	GACATCGATGTCCTCCCCAT (SEQ ID NO: 5590)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1577	PE2	PE4	RNF2	RNF2	+1 C to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5364)	AACGAACACCTCATGTAATGACTAAGATG (SEQ ID NO: 5365)	0	0	0	0
53C	HEK293T	Chen2021_1578	PEmax	PE4max	RNF2	RNF2	+1 C to A		GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5364)	AACGAACACCTCATGTAATGACTAAGATG (SEQ ID NO: 5365)	0	0	0	0
53C	HEK293T	Chen2021_1579	PE2	PE5	RNF2	RNF2	+1 C to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5364)	AACGAACACCTCATGTAATGACTAAGATG (SEQ ID NO: 5365)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1580	PEmax	PE5max	RNF2	RNF2	+1 C to A	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	ATGAGGTGTTCGTT (SEQ ID NO: 5364)	AACGAACACCTCATGTAATGACTAAGATG (SEQ ID NO: 5365)	0	0	GTCAACCATTAAGCAAAACAT (SEQ ID NO: 5269)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1581	PE2	PE4	FANCF	FANCF	+5 G to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	0	0
53C	HEK293T	Chen2021_1582	PEmax	PE4max	FANCF	FANCF	+5 G to T		GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	0	0
53C	HEK293T	Chen2021_1583	PE2	PE5	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1584	PEmax	PE5max	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1585	PE2	PE4	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGAGGAAGCGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGGAGGAAG (SEQ ID NO: 5243)	CGATTGCTTCAGACA (SEQ ID NO: 5259)	TGTCTGAAGCAATCGCTTCCTCCTGAAAAT (SEQ ID NO: 5260)	0	0	0	0

53C	HEK293T	Chen2021_1585	PEmax	PE4max	RUNX1	RUNX1	+5 G to T		GCATTTTCAGGGAGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGAGGAGGA (SEQ ID NO: 5243)	CGATTGCTTGAGAC (SEQ ID NO: 5269)	TGTCTGAAGCAATCGGCTTCCTCGCTGAAAAT (SEQ ID NO: 5280)	0	0	0	GATCAACCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1587	PE2	PE5	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGGAGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGAGGAGGA (SEQ ID NO: 5243)	CGATTGCTTGAGAC (SEQ ID NO: 5269)	TGTCTGAAGCAATCGGCTTCCTCGCTGAAAAT (SEQ ID NO: 5280)	0	0	0	GATCAACCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1588	PE2	PE5max	RUNX1	RUNX1	+5 G to T	+38	GCATTTTCAGGGAGGAAGGGA (SEQ ID NO: 5242)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ATTTTCAGAGGAGGA (SEQ ID NO: 5243)	CGATTGCTTGAGAC (SEQ ID NO: 5269)	TGTCTGAAGCAATCGGCTTCCTCGCTGAAAAT (SEQ ID NO: 5280)	0	0	0	GATCAACCA CTGTGGGTA CGA (SEQ ID NO: 5246)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1589	PE2	PE4	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGTGCCAAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	ATCAGTGCTCCAGAT (SEQ ID NO: 5616)	0	0	0	0	0
53C	HEK293T	Chen2021_1590	PE2	PEmax	DNMT1	DNMT1	+5 G to T		GATTCCTGGTGTGCCAAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	ATCAGTGCTCCAGAT (SEQ ID NO: 5616)	0	0	0	0	0
53C	HEK293T	Chen2021_1591	PE2	PE5	DNMT1	DNMT1	+5 G to T	+48	GATTCCTGGTGTGCCAAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	ATCAGTGCTCCAGAT (SEQ ID NO: 5616)	0	0	0	GCCCTTCAG CTAAAATAAA GG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1592	PE2	PE5max	DNMT1	DNMT1	+5 G to T	+48	GATTCCTGGTGTGCCAAGAAACA (SEQ ID NO: 5261)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCTGGTGCCAGA (SEQ ID NO: 5615)	ACAGTGGTGAC (SEQ ID NO: 5420)	ATCAGTGCTCCAGAT (SEQ ID NO: 5616)	0	0	0	GCCCTTCAG CTAAAATAAA GG (SEQ ID NO: 5270)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1593	PE2	PE4	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCAGTGCAGATGA (SEQ ID NO: 5285)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGAT (SEQ ID NO: 5440)	AATGTGCCATCTGGCAGCCTGCAG (SEQ ID NO: 5619)	0	0	0	0	0
53C	HEK293T	Chen2021_1594	PE2	PEmax	VEGFA	VEGFA	+5 G to T		GATGTCTGCAGGCAGTGCAGATGA (SEQ ID NO: 5285)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGAT (SEQ ID NO: 5440)	AATGTGCCATCTGGCAGCCTGCAG (SEQ ID NO: 5619)	0	0	0	0	0
53C	HEK293T	Chen2021_1596	PE2	PE5	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCAGTGCAGATGA (SEQ ID NO: 5285)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGAT (SEQ ID NO: 5440)	AATGTGCCATCTGGCAGCCTGCAG (SEQ ID NO: 5619)	0	0	0	GATGTACAAG AGAGCCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53C	HEK293T	Chen2021_1598	PE2	PE5max	VEGFA	VEGFA	+5 G to T	+57	GATGTCTGCAGGCAGTGCAGATGA (SEQ ID NO: 5285)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CTGCAGGCCAGA (SEQ ID NO: 5618)	TGAGTGCTCCAGAT (SEQ ID NO: 5440)	AATGTGCCATCTGGCAGCCTGCAG (SEQ ID NO: 5619)	0	0	0	GATGTACAAG AGAGCCCAG GGC (SEQ ID NO: 5271)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1597	PE2-P2A-BSD	PE2	HBB	HBB	+5 G to A		GCATTGGTGCACCTGCTGTGGTGC (SEQ ID NO: 5921)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTGTT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGGCGGTTCTATC TTAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5625)	0	0
52D	Hela	Chen2021_1598	PE2max-P2A-BSD	PE2max	HBB	HBB	+5 G to A		GCATTGGTGCACCTGCTGTGGTGC (SEQ ID NO: 5921)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTGTT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGGCGGTTCTATC TTAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5625)	0	0
52D	Hela	Chen2021_1599	PEmax-P2A-BSD	PEmax	HBB	HBB	+5 G to A		GCATTGGTGCACCTGCTGTGGTGC (SEQ ID NO: 5921)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTGTT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGGCGGTTCTATC TTAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5625)	0	0

52D	Hela	Chen2021_1609	PE2-P2A-BSD	PE3	HBB	HBB	+5 G to A	+72	GCATGGTGCAC CTGACTCCTG (SEQ ID NO: 5231)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5522)	CTGAAGAGAAGTGT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGT CAAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5814)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	GCCTTGATA CCAACCTGC CCA (SEQ ID NO: 5638)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1631	PEmax K-P2A-BSD	PE3max	HBB	HBB	+5 G to A	+72	GCATGGTGCAC CTGACTCCTG (SEQ ID NO: 5231)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5522)	CTGAAGAGAAGTGT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGT CAAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5814)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	GCCTTGATA CCAACCTGC CCA (SEQ ID NO: 5638)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1632	PEmax K-P2A-BSD	PE3max	HBB	HBB	+5 G to A	+72	GCATGGTGCAC CTGACTCCTG (SEQ ID NO: 5231)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5522)	CTGAAGAGAAGTGT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGT CAAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5814)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	GCCTTGATA CCAACCTGC CCA (SEQ ID NO: 5638)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1633	PE2-P2A-BSD	PE2	PRNP	PRNP	+6 G to T		GCAGTGGTGG GGGAGCCTTG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGCGGGGGGCCT (SEQ ID NO: 5627)	TTGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCGAAGGCC CCCCACCC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1634	PEmax K-P2A-BSD	PE2max	PRNP	PRNP	+6 G to T		GCAGTGGTGG GGGAGCCTTG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGCGGGGGGCCT (SEQ ID NO: 5627)	TTGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCGAAGGCC CCCCACCC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1635	PEmax K-P2A-BSD	PE4max	PRNP	PRNP	+6 G to T		GCAGTGGTGG GGGAGCCTTG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGCGGGGGGCCT (SEQ ID NO: 5627)	TTGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCGAAGGCC CCCCACCC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	0	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1636	PE2-P2A-BSD	PE3	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGG GGGAGCCTTG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGCGGGGGGCCT (SEQ ID NO: 5627)	TTGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCGAAGGCC CCCCACCC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	GCATGTTTT CAGGATAGT TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1637	PEmax K-P2A-BSD	PE3max	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGG GGGAGCCTTG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGCGGGGGGCCT (SEQ ID NO: 5627)	TTGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCGAAGGCC CCCCACCC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	GCATGTTTT CAGGATAGT TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1638	PEmax K-P2A-BSD	PE5max	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGG GGGAGCCTTG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGCGGGGGGCCT (SEQ ID NO: 5627)	TTGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCGAAGGCC CCCCACCC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC TAGTTACGCGTT AAACCAACTAGA A (SEQ ID NO: 5525)	GCATGTTTT CAGGATAGT TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1639	PE2-P2A-BSD	PE2	CDKL5	CDKL5	+1 C to T		GAGGGAGCTCCT AGAGGAGCTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGT CAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCC AATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGGTCAG GAGCCCCCCCCC CTGCACCCAGG GTCGGGGGGGCA ACCC (SEQ ID NO: 5638)	0	
52D	Hela	Chen2021_1640	PEmax K-P2A-BSD	PE2max	CDKL5	CDKL5	+1 C to T		GAGGGAGCTCCT AGAGGAGCTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGT CAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCC AATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGGTCAG GAGCCCCCCCCC CTGCACCCAGG GTCGGGGGGGCA ACCC (SEQ ID NO: 5638)	0	
52D	Hela	Chen2021_1641	PEmax K-P2A-BSD	PE4max	CDKL5	CDKL5	+1 C to T		GAGGGAGCTCCT AGAGGAGCTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA ATAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGT CAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCC AATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGGTCAG GAGCCCCCCCCC CTGCACCCAGG GTCGGGGGGGCA ACCC (SEQ ID NO: 5638)	0	

52D	Hela	Chen2021_1612	PE2–P2A–BSD	PE3	CDKL5	CDKL5	+1 C to T	+86	GAGGGACTCCTAGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCTGCACCCAGGAAAACCCTCAAAGTCGGGGGGCAACCC (SEQ ID NO: 5638)	GCAGAACCGCCACTCATTCA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1613	PEmax–P2A–BSD	PE3max	CDKL5	CDKL5	+1 C to T	+86	GAGGGACTCCTAGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCTGCACCCAGGAAAACCCTCAAAGTCGGGGGGCAACCC (SEQ ID NO: 5638)	GCAGAACCGCCACTCATTCA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1614	PEmax–P2A–BSD	PE5max	CDKL5	CDKL5	+1 C to T	+86	GAGGGACTCCTAGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCTGCACCCAGGAAAACCCTCAAAGTCGGGGGGCAACCC (SEQ ID NO: 5638)	GCAGAACCGCCACTCATTCA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1615	PE2–P2A–BSD	PE2	CXCR4	CXCR4	+5 G to C		GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
52D	Hela	Chen2021_1616	PEmax–P2A–BSD	PE2max	CXCR4	CXCR4	+5 G to C		GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
52D	Hela	Chen2021_1617	PEmax–P2A–BSD	PE4max	CXCR4	CXCR4	+5 G to C		GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
52D	Hela	Chen2021_1618	PE2–P2A–BSD	PE3	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	GCATCTTTGCCAACGTCAGTG (SEQ ID NO: 5646)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1619	PEmax–P2A–BSD	PE3max	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	GCATCTTTGCCAACGTCAGTG (SEQ ID NO: 5646)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1620	PEmax–P2A–BSD	PE5max	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	GCATCTTTGCCAACGTCAGTG (SEQ ID NO: 5646)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1621	PE2–P2A–BSD	PE2	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
52D	Hela	Chen2021_1622	PEmax–P2A–BSD	PE2max	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
52D	Hela	Chen2021_1623	PEmax–P2A–BSD	PE4max	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
52D	Hela	Chen2021_1624	PE2–P2A–BSD	PE3	IL2RB	IL2RB	+1 T to A, +5 G to C	+55	GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	GCTCCCTCCAAGTTGTCCACG (SEQ ID NO: 5651)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

52D	Hela	Chen2021_1625	PEmax-P2A-BSD	PE3max	IL2RB	IL2RB	+1 T to A, +5 G to C	+55	GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	GCTCCCTCCAAGTTGTCCACG (SEQ ID NO: 5651)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1626	PEmax-P2A-BSD	PE5max	IL2RB	IL2RB	+1 T to A, +5 G to C	+55	GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	GCTCCCTCCAAGTTGTCCACG (SEQ ID NO: 5651)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1627	PE2-P2A-BSD	PE2	HBG	HBG	+6 G to A, +8 C to A		GTTTGCCTTGTCAAGGCTAT (SEQ ID NO: 5648)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCTTGTCAAGGC (SEQ ID NO: 5649)	TATTGATAAAGGCAAGGCTGGCCA (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTATCAATAGCCTTGACAAGGCAA (SEQ ID NO: 5652)	0	0	0	0
52D	Hela	Chen2021_1628	PEmax-P2A-BSD	PE2max	HBG	HBG	+6 G to A, +8 C to A		GTTTGCCTTGTCAAGGCTAT (SEQ ID NO: 5648)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCTTGTCAAGGC (SEQ ID NO: 5649)	TATTGATAAAGGCAAGGCTGGCCA (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTATCAATAGCCTTGACAAGGCAA (SEQ ID NO: 5652)	0	0	0	0
52D	Hela	Chen2021_1629	PEmax-P2A-BSD	PE4max	HBG	HBG	+6 G to A, +8 C to A		GTTTGCCTTGTCAAGGCTAT (SEQ ID NO: 5648)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCTTGTCAAGGC (SEQ ID NO: 5649)	TATTGATAAAGGCAAGGCTGGCCA (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTATCAATAGCCTTGACAAGGCAA (SEQ ID NO: 5652)	0	0	0	0
52D	Hela	Chen2021_1630	PE2-P2A-BSD	PE3	HBG	HBG	+6 G to A, +8 C to A	+31	GTTTGCCTTGTCAAGGCTAT (SEQ ID NO: 5648)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCTTGTCAAGGC (SEQ ID NO: 5649)	TATTGATAAAGGCAAGGCTGGCCA (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTATCAATAGCCTTGACAAGGCAA (SEQ ID NO: 5652)	0	0	GCCTGGCTAAACTCCACCCAT (SEQ ID NO: 5653)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1631	PEmax-P2A-BSD	PE3max	HBG	HBG	+6 G to A, +8 C to A	+31	GTTTGCCTTGTCAAGGCTAT (SEQ ID NO: 5648)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCTTGTCAAGGC (SEQ ID NO: 5649)	TATTGATAAAGGCAAGGCTGGCCA (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTATCAATAGCCTTGACAAGGCAA (SEQ ID NO: 5652)	0	0	GCCTGGCTAAACTCCACCCAT (SEQ ID NO: 5653)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	Hela	Chen2021_1632	PEmax-P2A-BSD	PE5max	HBG	HBG	+6 G to A, +8 C to A	+31	GTTTGCCTTGTCAAGGCTAT (SEQ ID NO: 5648)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCTTGTCAAGGC (SEQ ID NO: 5649)	TATTGATAAAGGCAAGGCTGGCCA (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTATCAATAGCCTTGACAAGGCAA (SEQ ID NO: 5652)	0	0	GCCTGGCTAAACTCCACCCAT (SEQ ID NO: 5653)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1633	PE2	PE2	HBB	HBB	+5 G to A		GCATGGTGCACCTGACTCCTG (SEQ ID NO: 5621)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTCT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA (SEQ ID NO: 5625)	0	0
52D	HEK293T	Chen2021_1634	PEmax	PE2max	HBB	HBB	+5 G to A		GCATGGTGCACCTGACTCCTG (SEQ ID NO: 5621)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTCT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA (SEQ ID NO: 5625)	0	0
52D	HEK293T	Chen2021_1635	PEmax	PE4max	HBB	HBB	+5 G to A		GCATGGTGCACCTGACTCCTG (SEQ ID NO: 5621)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTCT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA (SEQ ID NO: 5625)	0	0
52D	HEK293T	Chen2021_1636	PE2	PE3	HBB	HBB	+5 G to A	+72	GCATGGTGCACCTGACTCCTG (SEQ ID NO: 5621)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTCT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA (SEQ ID NO: 5625)	GCCTTGATACCAACCTGCCCA (SEQ ID NO: 5630)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1637	PEmax	PE3max	HBB	HBB	+5 G to A	+72	GCATGGTGCACCTGACTCCTG (SEQ ID NO: 5621)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTCT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA (SEQ ID NO: 5625)	GCCTTGATACCAACCTGCCCA (SEQ ID NO: 5630)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1638	PEmax	PE5max	HBB	HBB	+5 G to A	+72	GCATGGTGCACCTGACTCCTG (SEQ ID NO: 5621)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGCACCTGACTC (SEQ ID NO: 5622)	CTGAAGAGAAGTCT (SEQ ID NO: 5623)	AGACTTCTCTTCAGGAGTCAGGTGCAC (SEQ ID NO: 5624)	AACTCTCT (SEQ ID NO: 5614)	CGCGGTTCTATCTAGTTACGCGTTAAACCAACTAGAA (SEQ ID NO: 5625)	GCCTTGATACCAACCTGCCCA (SEQ ID NO: 5630)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

52D	HEK293T	Chen2021_1639	PE2	PRNP	PRNP	+6 G to T		GCAGTGGTGG GCGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGGCCAAGGCC CCCCACC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC GTAGTTATCGCGTT AAACCCAACTAGA (SEQ ID NO: 5525)	0				
52D	HEK293T	Chen2021_1640	PEmax	PRNP	PRNP	+6 G to T		GCAGTGGTGG GCGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGGCCAAGGCC CCCCACC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC GTAGTTATCGCGTT AAACCCAACTAGA (SEQ ID NO: 5525)	0				
52D	HEK293T	Chen2021_1641	PEmax	PRNP	PRNP	+6 G to T		GCAGTGGTGG GCGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGGCCAAGGCC CCCCACC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC GTAGTTATCGCGTT AAACCCAACTAGA (SEQ ID NO: 5525)	0				
52D	HEK293T	Chen2021_1642	PE2	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGG GCGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGGCCAAGGCC CCCCACC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC GTAGTTATCGCGTT AAACCAACTAGA (SEQ ID NO: 5525)	GCATGTTTT CACGATAGT TAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)			
52D	HEK293T	Chen2021_1643	PEmax	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGG GCGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGGCCAAGGCC CCCCACC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC GTAGTTATCGCGTT AAACCAACTAGA (SEQ ID NO: 5525)	GCATGTTTT CACGATAGT TAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)			
52D	HEK293T	Chen2021_1644	PEmax	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGG GCGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGGCCAAGGCC CCCCACC (SEQ ID NO: 5629)	TAGACACA (SEQ ID NO: 5631)	CGCGGTTCTATC GTAGTTATCGCGTT AAACCAACTAGA (SEQ ID NO: 5525)	GCATGTTTT CACGATAGT TAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)			
52D	HEK293T	Chen2021_1645	PE2	CDKL5	CDKL5	+1 C to T		GAGGGACTCCT AGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GAGTGGGAATTGTGT (SEQ ID NO: 5634)	ATTGTGACACAATTCCCCC AATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAG GAGCCCCCCCC CTGCACCCAGG AAAACCCTCAAA GTCGGGGGGCA ACCC (SEQ ID NO: 5638)	0				
52D	HEK293T	Chen2021_1646	PEmax	CDKL5	CDKL5	+1 C to T		GAGGGACTCCT AGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GAGTGGGAATTGTGT (SEQ ID NO: 5634)	ATTGTGACACAATTCCCCC AATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAG GAGCCCCCCCC CTGCACCCAGG AAAACCCTCAAA GTCGGGGGGCA ACCC (SEQ ID NO: 5638)	0				
52D	HEK293T	Chen2021_1647	PEmax	CDKL5	CDKL5	+1 C to T		CAGGGACTCCT AGAGGACTG (SEQ ID NO: 5633)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTGCTAGAGGGA (SEQ ID NO: 5633)	TTGTGGGAATTGTGT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCC AATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAG GAGCCCCCCCC CTGCACCCAGG AAAACCCTCAAA GTCGGGGGGCA ACCC (SEQ ID NO: 5638)	0				
52D	HEK293T	Chen2021_1648	PE2	PE3	CDKL5	+1 C to T	+86	GAGGGACTCCT AGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTGCTAGAGGGA (SEQ ID NO: 5633)	TTGTGGGAATTGTGT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCC AATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAG GAGCCCCCCCC CTGCACCCAGG AAAACCCTCAAA GTCGGGGGGCA ACCC (SEQ ID NO: 5638)	GCAGAACCG CCACTCATT CA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)			
52D	HEK293T	Chen2021_1649	PEmax	CDKL5	CDKL5	+1 C to T	+86	GAGGGACTCCT AGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTGCTAGAGGGA (SEQ ID NO: 5633)	TTGTGGGAATTGTGT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCC AATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAG GAGCCCCCCCC CTGCACCCAGG AAAACCCTCAAA GTCGGGGGGCA ACCC (SEQ ID NO: 5638)	GCAGAACCG CCACTCATT CA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)			

52D	HEK293T	Chen2021_1650	PEmax	PE5max	CCR5	CCRL5	+1 C to T	+86	GAGGGACTCCTTAGAGGTAG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	ACCACCCACAAGT (SEQ ID NO: 5640)	TTGGCGTAGAAGCG (GTCA (SEQ ID NO: 5641)	ATATTGACAGCAATTCGTGT (AATCGTCTAGGAGTC (SEQ ID NO: 5634)	AAAAAGCAC (SEQ ID NO: 5637)	GTCAGGGTCAAG (GAGGCCCCCCC (CTGCAGCCAGG (AAAACCCTCAAA (GTCGGGGGGCA (ACCC (SEQ ID NO: 5638)	GCAGAACCG (CCACTCATT (CA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1651	PE2	CXCR4	CXCR4	CXCR4	+5 G to C		GCAACCACCCA (CAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGT (CA (SEQ ID NO: 5640)	TTGGCGTAGAAGCG (GTCA (SEQ ID NO: 5641)	TGAGCGCTTGTACGCCAAG (TGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
52D	HEK293T	Chen2021_1652	PEmax	PE2max	CXCR4	CXCR4	+5 G to C		GCAACCACCCA (CAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGT (CA (SEQ ID NO: 5640)	TTGGCGTAGAAGCG (GTCA (SEQ ID NO: 5641)	TGAGCGCTTGTACGCCAAG (TGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
52D	HEK293T	Chen2021_1653	PEmax	PE4max	CXCR4	CXCR4	+5 G to C		GCAACCACCCA (CAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGT (CA (SEQ ID NO: 5640)	TTGGCGTAGAAGCG (GTCA (SEQ ID NO: 5641)	TGAGCGCTTGTACGCCAAG (TGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
52D	HEK293T	Chen2021_1654	PE2	PE3	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCA (CAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGT (CA (SEQ ID NO: 5640)	TTGGCGTAGAAGCG (GTCA (SEQ ID NO: 5641)	TGAGCGCTTGTACGCCAAG (TGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	GCATCTTTG (CCAACGTCA (GTG (SEQ ID NO: 5645)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1655	PEmax	PE3max	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCA (CAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGT (CA (SEQ ID NO: 5640)	TTGGCGTAGAAGCG (GTCA (SEQ ID NO: 5641)	TGAGCGCTTGTACGCCAAG (TGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	GCATCTTTG (CCAACGTCA (GTG (SEQ ID NO: 5645)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1656	PEmax	PE5max	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCA (CAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGT (CA (SEQ ID NO: 5640)	TTGGCGTAGAAGCG (GTCA (SEQ ID NO: 5641)	TGAGCGCTTGTACGCCAAG (TGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	GCATCTTTG (CCAACGTCA (GTG (SEQ ID NO: 5645)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1657	PE2	IL2RB	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCT (TTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5646)	AAGTCGGAGGCTTG (GGA (SEQ ID NO: 441)	TCCCAAGCGTCCGACTTCG (TTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
52D	HEK293T	Chen2021_1658	PEmax	PE2max	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCT (TTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5646)	AAGTCGGAGGCTTG (GGA (SEQ ID NO: 441)	TCCCAAGCGTCCGACTTCG (TTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
52D	HEK293T	Chen2021_1659	PEmax	PE4max	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCT (TTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5646)	AAGTCGGAGGCTTG (GGA (SEQ ID NO: 441)	TCCCAAGCGTCCGACTTCG (TTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
52D	HEK293T	Chen2021_1660	PE2	PE3	IL2RB	IL2RB	+1 T to A, +5 G to C	+55	GCCAGGTGTCT (TTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5646)	AAGTCGGAGGCTTG (GGA (SEQ ID NO: 441)	TCCCAAGCGTCCGACTTCG (TTTGAAAGACAC (SEQ ID NO: 5647)	0	0	GCTCCGTCC (AAGTTGTCC (ACG (SEQ ID NO: 5651)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1661	PEmax	PE3max	IL2RB	IL2RB	+1 T to A, +5 G to C	+55	GCCAGGTGTCT (TTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5646)	AAGTCGGAGGCTTG (GGA (SEQ ID NO: 441)	TCCCAAGCGTCCGACTTCG (TTTGAAAGACAC (SEQ ID NO: 5647)	0	0	GCTCCGTCC (AAGTTGTCC (ACG (SEQ ID NO: 5651)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1662	PEmax	PE5max	IL2RB	IL2RB	+1 T to A, +5 G to C	+55	GCCAGGTGTCT (TTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5646)	AAGTCGGAGGCTTG (GGA (SEQ ID NO: 441)	TCCCAAGCGTCCGACTTCG (TTTGAAAGACAC (SEQ ID NO: 5647)	0	0	GCTCCGTCC (AAGTTGTCC (ACG (SEQ ID NO: 5651)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1663	PE2	HBG	HBG	HBG	+5 G to A, +9 C to A		GTTTGCCTTGT (CAAGGGTAT (SEQ ID NO: 5648)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCTTGTGTAG (GTGCAACCAG (SEQ ID NO: 5649)	TATTGATAAAGGCA (AGGGTGGCCA (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTAG (TCAATAGCCTTGACAAGGC (CAA (SEQ ID NO: 5652)	0	0	0	0
52D	HEK293T	Chen2021_1664	PEmax	PE2max	HBG	HBG	+5 G to A, +9 C to A		GTTTGCCTTGT (CAAGGGTAT (SEQ ID NO: 5648)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TTGCCTTGTGTAG (GTGCAACCAG (SEQ ID NO: 5649)	TATTGATAAAGGCA (AGGGTGGCCA (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTAG (TCAATAGCCTTGACAAGGC (CAA (SEQ ID NO: 5652)	0	0	0	0

52D	HEK293T	Chen2021_1665	PEmax	PE4max	HBG	HBG	+6 G to A, +8 C to A	GTTTTGCCTTGTCCAAGCGCTATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5645)	TTGCCTTGTCCAAGCGCTATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5646)	TATTGATAAAGGCAGCTGGCCATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTAAGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5652)	0	0	0	GCCTGGCCTAAGCTCAGCCCAT (SEQ ID NO: 5653)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1666	PE2	PE3	HBG	HBG	+6 G to A, +8 C to A	GTTTTGCCTTGTCCAAGCGCTATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5649)	TTGCCTTGTCCAAGCGCTATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5649)	TATTGATAAAGGCAGCTGGCCATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTAAGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5652)	0	0	0	GCCTGGCCTAAGCTCAGCCCAT (SEQ ID NO: 5653)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1667	PEmax	PE3max	HBG	HBG	+6 G to A, +8 C to A	GTTTTGCCTTGTCCAAGCGCTATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5649)	TTGCCTTGTCCAAGCGCTATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5649)	TATTGATAAAGGCAGCTGGCCATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTAAGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5652)	0	0	0	GCCTGGCCTAAGCTCAGCCCAT (SEQ ID NO: 5653)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52D	HEK293T	Chen2021_1668	PEmax	PE5max	HBG	HBG	+6 G to A, +8 C to A	GTTTTGCCTTGTCCAAGCGCTATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5649)	TTGCCTTGTCCAAGCGCTATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5649)	TATTGATAAAGGCAGCTGGCCATTAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5650)	TGGCCAGCCTTGCCTTTAAGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5652)	0	0	0	GCCTGGCCTAAGCTCAGCCCAT (SEQ ID NO: 5653)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52E.53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1669	mock	Control	CDKL5							0	0	0	0	0	0	0	0	0	0	0
52E.53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1670	PE2	PE2 (MLH1dn 0 ug)	CDKL5	CDKL5	+1 C to T, +13 T ins	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGATGTTGTGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	0	GTCAGGGTCAGGAGCCCCCCCCCTCTGCACCCAGGGAAAACCCTCAAAGTCGGGGCGGCAACCC (SEQ ID NO: 5638)	0	0	0	0	0
52E.53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1671	PE2	PE4 (MLH1dn 1 ug)	CDKL5	CDKL5	+1 C to T, +13 T ins	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGATGTTGTGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	0	GTCAGGGTCAGGAGCCCCCCCCCTCTGCACCCAGGGAAAACCCTCAAAGTCGGGGCGGCAACCC (SEQ ID NO: 5638)	0	0	0	0	0
52E.53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1672	PE2	PE4 (MLH1dn 2 ug)	CDKL5	CDKL5	+1 C to T, +13 T ins	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGATGTTGTGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	0	GTCAGGGTCAGGAGCCCCCCCCCTCTGCACCCAGGGAAAACCCTCAAAGTCGGGGCGGCAACCC (SEQ ID NO: 5638)	0	0	0	0	0
52E.53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1673	PE2	PE3 (MLH1dn 0 ug)	CDKL5	CDKL5	+1 C to T, +13 T ins	+86	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGATGTTGTGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCCTCTGCACCCAGGGAAAACCCTCAAAGTCGGGGCGGCAACCC (SEQ ID NO: 5638)	GCAGAACCGTCCAACTCATTCA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1674	PE2	PE5 (MLH1dn 1 ug)	CDKL5	CDKL5	+1 C to T, +13 T ins	+86	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGATGTTGTGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCCTCTGCACCCAGGGAAAACCCTCAAAGTCGGGGCGGCAACCC (SEQ ID NO: 5638)	GCAGAACCGTCCAACTCATTCA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52E.53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1675	PE2	PE5 (MLH1dn 2 ug)	CDKL5	CDKL5	+1 C to T, +13 T ins	+86	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGATGTTGTGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCCTCTGCACCCAGGGAAAACCCTCAAAGTCGGGGCGGCAACCC (SEQ ID NO: 5638)	GCAGAACCGTCCAACTCATTCA (SEQ ID NO: 5643)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

52E,53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1676	PE2	PE3b (MLH1dn 0 µg)	CDKL5	CDKL5	+1 C to T, +13 T ins	-3	GAGGGACTCCTAGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCTGCACCCAGGAAAACCCTCAAAGTCGGGGGGCAACCC (SEQ ID NO: 5638)	ACACAATTCCCCAATCCTCT (SEQ ID NO: 5654)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1677	PE2	PE5b (MLH1dn 1 µg)	CDKL5	CDKL5	+1 C to T, +13 T ins	-3	GAGGGACTCCTAGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCTGCACCCAGGAAAACCCTCAAAGTCGGGGGGCAACCC (SEQ ID NO: 5638)	ACACAATTCCCCAATCCTCT (SEQ ID NO: 5654)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52E,53D	Human iPSC with CDKL5 c.1412delA mutation	Chen2021_1678	PE2	PE5b (MLH1dn 2 µg)	CDKL5	CDKL5	+1 C to T, +13 T ins	-3	GAGGGACTCCTAGAGGACTG (SEQ ID NO: 5632)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GACTCCTAGAGGA (SEQ ID NO: 5633)	TTGGGGAATTGTGTCAATAT (SEQ ID NO: 5634)	ATATTGACACAATTCCCCAATCCTCTAGGAGTC (SEQ ID NO: 5635)	AAAAACAC (SEQ ID NO: 5637)	GTCAGGGTCAGGAGCCCCCCCCTGCACCCAGGAAAACCCTCAAAGTCGGGGGGCAACCC (SEQ ID NO: 5638)	ACACAATTCCCCAATCCTCT (SEQ ID NO: 5654)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1679	mock	Control	PRNP								0	0	0	0	0
52F,53E	Human T cells	Chen2021_1680	PE2	PE3 (MLH1dn 0 µg)	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGGGGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCCAAGGCCCCCCACC (SEQ ID NO: 5629)	0	0	GCATGTTTTCACGATAGTAA (SEQ ID NO: 5636)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1681	PE2	PE5 (MLH1dn 0.5 µg)	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGGGGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCCAAGGCCCCCCACC (SEQ ID NO: 5629)	0	0	GCATGTTTTCACGATAGTAA (SEQ ID NO: 5636)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1682	PE2	PE5 (MLH1dn 1 µg)	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGGGGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCCAAGGCCCCCCACC (SEQ ID NO: 5629)	0	0	GCATGTTTTCACGATAGTAA (SEQ ID NO: 5636)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52F,53E	Human T cells	Chen2021_1683	PE2	PE5 (MLH1dn 2 µg)	PRNP	PRNP	+6 G to T	+72	GCAGTGGTGGGGGGCCTTGG (SEQ ID NO: 5626)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GGTGGGGGGCCT (SEQ ID NO: 5627)	TGGCGTCTACAT (SEQ ID NO: 5628)	ATGTAGACGCCAAGGCCCCCCACC (SEQ ID NO: 5629)	0	0	GCATGTTTTCACGATAGTAA (SEQ ID NO: 5636)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1684	mock	Control	FANCF								0	0	0	0	0
52F,53E	Human T cells	Chen2021_1685	PE2	PE3 (MLH1dn 0 µg)	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1686	PE2	PE5 (MLH1dn 0.5 µg)	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1687	PE2	PE5 (MLH1dn 1 µg)	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52F,53E	Human T cells	Chen2021_1688	PE2	PE5 (MLH1dn 2 µg)	FANCF	FANCF	+5 G to T	+48	GGAATCCCTTCTGCAGCACC (SEQ ID NO: 5255)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	TCCCTTCTGCAGC (SEQ ID NO: 5610)	ACCTTGATCGCTTTTCC (SEQ ID NO: 5611)	GGAAAAGCGATCAAGGTGCTGCAGAAGGGA (SEQ ID NO: 5612)	0	0	GGGGTCCCAGGTGCTGACGT (SEQ ID NO: 924)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1689	mock	Control	RNF2								0	0	0	0	0

52F,53E	Human T cells	Chen2021_1690	PE2	PE3 (MLH1dn 0 μg)	RNF2	RNF2	+1 T ins	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TCTGAGGTGTTCGTT (SEQ ID NO: 5655)	AACGAACACCTCAGAGTAATGACTAAGATG (SEQ ID NO: 5656)	0	0	TCAACCATTAAGCAAAACAT (SEQ ID NO: 5657)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1691	PE2	PE5 (MLH1dn 0.5 μg)	RNF2	RNF2	+1 T ins	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TCTGAGGTGTTCGTT (SEQ ID NO: 5655)	AACGAACACCTCAGAGTAATGACTAAGATG (SEQ ID NO: 5656)	0	0	TCAACCATTAAGCAAAACAT (SEQ ID NO: 5657)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1692	PE2	PE5 (MLH1dn 1 μg)	RNF2	RNF2	+1 T ins	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TCTGAGGTGTTCGTT (SEQ ID NO: 5655)	AACGAACACCTCAGAGTAATGACTAAGATG (SEQ ID NO: 5656)	0	0	TCAACCATTAAGCAAAACAT (SEQ ID NO: 5657)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52F,53E	Human T cells	Chen2021_1693	PE2	PE5 (MLH1dn 2 μg)	RNF2	RNF2	+1 T ins	+41	GTCATCTTAGTCATTACCTG (SEQ ID NO: 5251)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	CATCTTAGTCATTAC (SEQ ID NO: 5252)	TCTGAGGTGTTCGTT (SEQ ID NO: 5655)	AACGAACACCTCAGAGTAATGACTAAGATG (SEQ ID NO: 5656)	0	0	TCAACCATTAAGCAAAACAT (SEQ ID NO: 5657)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1694	mock	Control	CXCR4								0	0	0	0	0
52F,53E	Human T cells	Chen2021_1695	PE2	PE2 (MLH1dn 0 μg)	CXCR4	CXCR4	+5 G to C		GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
53E	Human T cells	Chen2021_1696	PE2	PE4 (MLH1dn 1 μg)	CXCR4	CXCR4	+5 G to C		GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
52F,53E	Human T cells	Chen2021_1697	PE2	PE4 (MLH1dn 2 μg)	CXCR4	CXCR4	+5 G to C		GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	0	0
52F,53E	Human T cells	Chen2021_1698	PE2	PE3 (MLH1dn 0 μg)	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	CATCTTTGCCAACGTCAGTG (SEQ ID NO: 5658)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1699	PE2	PE5 (MLH1dn 1 μg)	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	CATCTTTGCCAACGTCAGTG (SEQ ID NO: 5658)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
52F,53E	Human T cells	Chen2021_1700	PE2	PE5 (MLH1dn 2 μg)	CXCR4	CXCR4	+5 G to C	+43	GCAACCACCCACAAGTCATTG (SEQ ID NO: 5639)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	ACCACCCACAAGTCA (SEQ ID NO: 5640)	TTGGCGTAGAAGCGGTCA (SEQ ID NO: 5641)	TGACCGCTTCTACGCCAATGACTTGTGGGTGGT (SEQ ID NO: 5642)	0	0	CATCTTTGCCAACGTCAGTG (SEQ ID NO: 5658)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)
53E	Human T cells	Chen2021_1701	mock	Control	IL2RB								0	0	0	0	0
52F,53E	Human T cells	Chen2021_1702	PE2	PE2 (MLH1dn 0 μg)	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
53E	Human T cells	Chen2021_1703	PE2	PE4 (MLH1dn 1 μg)	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
52F,53E	Human T cells	Chen2021_1704	PE2	PE4 (MLH1dn 2 μg)	IL2RB	IL2RB	+1 T to A, +5 G to C		GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	0	0
52F,53E	Human T cells	Chen2021_1705	PE2	PE3 (MLH1dn 0 μg)	IL2RB	IL2RB	+1 T to A, +5 G to C	+55	GCCAGGTGTCTTTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTGGGA (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCTTTGAAAGACAC (SEQ ID NO: 5647)	0	0	CTCCCTCCAAGTTGTCCACG (SEQ ID NO: 5659)	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)

53E	Human T cells	Chen2021.1735	PE2	PE5 (MLH1dn 1 ug)	IL2RB	+1 T to A +5 G to C	+55	GCCAGGTGTCT TTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTG G3A (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCG TTTGAAAGACAC (SEQ ID NO: 5647)	G		CTCCCTTCCA ACTTGTCCA CG (SEQ ID NO: 5659)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)
52F53E	Human T cells	Chen2021.1737	PE2	PE5 (MLH1dn 2 ug)	IL2RB	+1 T to A +5 G to C	+55	GCCAGGTGTCT TTCAAAGTAG (SEQ ID NO: 5644)	GTTTTAGAGCTAGAAATAGCAAGTTAAAA TAAGGCTAGTCCGTTATCAACTTGAAAAA GTGGCACCGAGTCGGTGC (SEQ ID NO: 5237)	GTGTCTTTCAAAG (SEQ ID NO: 5645)	AAGTCGGAGGCTTG G3A (SEQ ID NO: 441)	TCCCAAGCCTCCGACTTCG TTTGAAAGACAC (SEQ ID NO: 5647)	G		CTCCCTTCCA ACTTGTCCA CG (SEQ ID NO: 5659)	GTTTTAGAGCTAGAAATAGCAAG TTAAAATAAGGCTAGTCCGTTAT CAACTTGAAAAAGTGGCACCGA GTCGGTGC (SEQ ID NO: 5237)

Example 2: Development of PEmax

[791] To further improve prime editing, the PE2 protein was optimized by varying reverse transcriptase (RT) codon usages, the length and composition of the peptide linkers between nCas9 and the reverse transcriptase, the location, composition, and number of NLS sequences, and mutations within the SpCas9 domain (FIGs. 55A and 55B). Among 20 such variants tested, the greatest enhancement in editing efficiency was observed with a prime editor architecture that uses a Genscript human codon-optimized RT, a 34-aa linker containing a bipartite SV40 NLS (Wu *et al.*, 2009), an additional C-terminal c-Myc NLS (Dang and Lee, 1988), and R221K and N394K mutations in SpCas9 previously shown to improve Cas9 nuclease activity (Spencer and Zhang, 2017) (FIGs. 54A and 55A). This optimized prime editor architecture was designated as PEmax. Across seven substitution edits targeting different loci, using the PEmax architecture with the PE2 system (PE2max) increased the average frequency of intended editing by 2.3-fold in HeLa cells and 1.2-fold in HEK293T cells over the original PE2 architecture (FIG. 55B). Similarly, PE3 using the PEmax architecture (PE3max) increased average editing efficiencies over PE3 by 3.2-fold in HeLa cells and 1.2-fold in HEK293T cells, without substantially changing product purity (FIGs. 54A and 55A).

REFERENCES

[792] Acharya, S., Wilson, T., Gradia, S., Kane, M.F., Guerrette, S., Marsischky, G.T., Kolodner, R., and Fishel, R. (1996). hMSH2 forms specific mispair-binding complexes with hMSH3 and hMSH6. Proceedings of the National Academy of Sciences *93*, 13629-13634.

[793] Amato, A., Cappabianca, M.P., Perri, M., Zaghis, I., Grisanti, P., Ponzini, D., and Di Biagio, P. (2014). Interpreting elevated fetal hemoglobin in pathology and health at the basic laboratory level: new and known γ- gene mutations associated with hereditary persistence of fetal hemoglobin. International Journal of Laboratory Hematology *36*, 13-19.

[794] Anzalone, A.V., Koblan, L.W., and Liu, D.R. (2020). Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. Nature Biotechnology *38*, 824-844.

[795] Anzalone, A.V., Randolph, P.B., Davis, J.R., Sousa, A.A., Koblan, L.W., Levy, J.M., Chen, P.J., Wilson, C., Newby, G.A., Raguram, A., *et al.* (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. Nature.

[796] Asante, E.A., Smidak, M., Grimshaw, A., Houghton, R., Tomlinson, A., Jeelani, A., Jakubcova, T., Hamdan, S., Richard-Londt, A., Linehan, J.M., *et al.* (2015). A naturally occurring variant of the human prion protein completely prevents prion disease. Nature *522*, 478-481.

[797] Bacher, J.W., Flanagan, L.A., Smalley, R.L., Nassif, N.A., Burgart, L.J., Halberg, R.B., Megid, W.M.A., and Thibodeau, S.N. (2004). Development of a fluorescent multiplex assay for detection of MSI-High tumors. Disease Markers *20*, 237-250.

[798] Bartlett, D.W., and Davis, M.E. (2006). Insights into the kinetics of siRNA-mediated gene silencing from live-cell and live-animal bioluminescent imaging. Nucleic Acids Research *34*, 322-333.

[799] Bosch, J.A., Birchak, G., and Perrimon, N. (2021). Precise genome engineering in Drosophila using prime editing. Proceedings of the National Academy of Sciences *118*, e2021996118.

[800] Bothmer, A., Phadke, T., Barrera, L.A., Margulies, C.M., Lee, C.S., Buquicchio, F., Moss, S., Abdulkerim, H.S., Selleck, W., Jayaram, H., *et al.* (2017). Characterization of the interplay between DNA repair and CRISPR/Cas9-induced DNA lesions at an endogenous locus. Nature Communications *8*, 13905.

[801] Cavaleiro, A.M., Kim, S.H., Seppälä, S., Nielsen, M.T., and Nørholm, M.H.H. (2015). Accurate DNA Assembly and Genome Engineering with Optimized Uracil Excision Cloning. ACS Synthetic Biology *4*, 1042-1046.

[802] Chen, P.-F., Chen, T., Forman, T.E., Swanson, A.C., O'Kelly, B., Dwyer, S.A., Buttermore, E.D., Kleiman, R., Js Carrington, S., Lavery, D.J., *et al.* (2021). Generation and characterization of human induced pluripotent stem cells (iPSCs) from three male and three female patients with CDKL5 Deficiency Disorder (CDD). Stem Cell Research *53*, 102276.

[803] Clement, K., Rees, H., Canver, M.C., Gehrke, J.M., Farouni, R., Hsu, J.Y., Cole, M.A., Liu, D.R., Joung, J.K., Bauer, D.E., *et al.* (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. Nature Biotechnology *37*, 224-226.

[804] Dang, C.V., and Lee, W.M. (1988). Identification of the human c-myc protein nuclear translocation signal. Molecular and Cellular Biology *8*, 4048-4054.

[805] Engler, C., Kandzia, R., and Marillonnet, S. (2008). A One Pot, One Step, Precision Cloning Method with High Throughput Capability. PLoS ONE *3*, e3647.

[806] Fang, W.H., and Modrich, P. (1993). Human strand-specific mismatch repair occurs by a bidirectional mechanism similar to that of the bacterial reaction. Journal of Biological Chemistry *268*, 11838-11844.

[807] Fishel, R., Lescoe, M.K., Rao, M.R.S., Copeland, N.G., Jenkins, N.A., Garber, J., Kane, M., and Kolodner, R. (1993). The human mutator gene homolog MSH2 and its association with hereditary nonpolyposis colon cancer. Cell *75*, 1027-1038.

[808] Geng, H., Du, C., Chen, S., Salerno, V., Manfredi, C., and Hsieh, P. (2011). In vitro studies of DNA mismatch repair proteins. Analytical Biochemistry *413*, 179-184.

[809] Genschel, J., Bazemore, L.R., and Modrich, P. (2002). Human Exonuclease I Is Required for 5′ and 3′ Mismatch Repair. Journal of Biological Chemistry *277*, 13302-13311.

[810] Genschel, J., Littman, S.J., Drummond, J.T., and Modrich, P. (1998). Isolation of MutSβ from Human Cells and Comparison of the Mismatch Repair Specificities of MutSβ and MutSα. Journal of Biological Chemistry *273*, 19895-19901.

[811] Gilbert, Luke A., Horlbeck, Max A., Adamson, B., Villalta, Jacqueline E., Chen, Y., Whitehead, Evan H., Guimaraes, C., Panning, B., Ploegh, Hidde L., Bassik, Michael C., *et al.* (2014). Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell *159*, 647-661.

[812] Gilbert, Luke A., Larson, Matthew H., Morsut, L., Liu, Z., Brar, Gloria A., Torres, Sandra E., Stern-Ginossar, N., Brandman, O., Whitehead, Evan H., Doudna, Jennifer A., *et al.* (2013). CRISPR-Mediated Modular RNA-Guided Regulation of Transcription in Eukaryotes. Cell *154*, 442-451.

[813] Gueneau, E., Dherin, C., Legrand, P., Tellier-Lebegue, C., Gilquin, B., Bonnesoeur, P., Londino, F., Quemener, C., Le Du, M.-H., Márquez, J.A., *et al.* (2013). Structure of the MutLα C-terminal domain reveals how Mlh1 contributes to Pms1 endonuclease site. Nature Structural & Molecular Biology *20*, 461-468.

[814] Guerrette, S., Acharya, S., and Fishel, R. (1999). The Interaction of the Human MutL Homologues in Hereditary Nonpolyposis Colon Cancer. Journal of Biological Chemistry *274*, 6336-6341.

[815] Gupta, S., Gellert, M., and Yang, W. (2012). Mechanism of mismatch recognition revealed by human MutSβ bound to unpaired DNA loops. Nature Structural & Molecular Biology *19*, 72-78.

[816] Hempelmann, J.A., Scroggins, S.M., Pritchard, C.C., and Salipante, S.J. (2015). MSIplus for Integrated Colorectal Cancer Molecular Testing by Next-Generation Sequencing. The Journal of Molecular Diagnostics *17*, 705-714.

[817] Hussmann, J.A., Ling, J., Ravisankar, P., Yan, J., Cirincione, A., Xu, A., Simpson, D., Yang, D., Bothmer, A., Cotta-Ramusino, C., *et al.* (2021). Repair-seq enables systematic mapping of DNA repair processes in genome editing. Submitted.

[818] Iaccarino, I., Marra, G., and Palomba, F.J., J. (1998). hMSH2 and hMSH6 play distinct roles in mismatch binding and contribute differently to the ATPase activity of hMutSalpha. The EMBO Journal *17*, 2677-2686.

[819] Ingram, V.M. (1956). A Specific Chemical Difference Between the Globins of Normal Human and Sickle-Cell Anæmia Hæmoglobin. Nature *178*, 792-794.

[820] Iyer, R.R., Pluciennik, A., Burdett, V., and Modrich, P.L. (2006). DNA Mismatch Repair: Functions and Mechanisms. Chemical Reviews *106*, 302-323.

[821] Jin, S., Lin, Q., Luo, Y., Zhu, Z., Liu, G., Li, Y., Chen, K., Qiu, J.-L., and Gao, C. (2021). Genome-wide specificity of prime editors in plants. Nature Biotechnology.

[822] Kadyrov, F.A., Dzantiev, L., Constantin, N., and Modrich, P. (2006). Endonucleolytic Function of MutLα in Human Mismatch Repair. Cell *126*, 297-308.

[823] Kim, D.Y., Moon, S.B., Ko, J.-H., Kim, Y.-S., and Kim, D. (2020). Unbiased investigation of specificities of prime editing systems in human cells. Nucleic Acids Research *48*, 10576-10589.

[824] Kim, J.H., Lee, S.-R., Li, L.-H., Park, H.-J., Park, J.-H., Lee, K.Y., Kim, M.-K., Shin, B.A., and Choi, S.-Y. (2011). High Cleavage Efficiency of a 2A Peptide Derived from Porcine Teschovirus-1 in Human Cell Lines, Zebrafish and Mice. PLoS ONE *6*, e18556.

[825] Kunkel, T.A., and Erie, D.A. (2005). DNA mismatch repair. Annu Rev Biochem *74*, 681-710.

[826] Lahue, R.S., Au, K.G., and Modrich, P. (1989). DNA mismatch correction in a defined system. Science *245*, 160.

[827] Leach, F.S., Nicolaides, N.C., Papadopoulos, N., Liu, B., Jen, J., Parsons, R., Peltomäki, P., Sistonen, P., Aaltonen, L.A., Nyström-Lahti, M., *et al.* (1993). Mutations of a mutS homolog in hereditary nonpolyposis colorectal cancer. Cell *75*, 1215-1225.

[828] Li, G.-M. (2008). Mechanisms and functions of DNA mismatch repair. Cell Research *18*, 85-98.

[829] Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., Wang, Y., Anzalone, A.V., Raguram, A., Doman, J.L., *et al.* (2020). Prime genome editing in rice and wheat. Nature Biotechnology *38*, 582-585.

[830] Liu, P., Liang, S.-Q., Zheng, C., Mintzer, E., Zhao, Y.G., Ponnienselvan, K., Mir, A., Sontheimer, E.J., Gao, G., Flotte, T.R., *et al.* (2021). Improved prime editors enable pathogenic allele correction and cancer modelling in adult mice. Nature Communications *12*.

[831] Liu, S., Wang, Q., Yu, X., Li, Y., Guo, Y., Liu, Z., Sun, F., Hou, W., Li, C., Wu, L., *et al.* (2018). HIV-1 inhibition in cells with CXCR4 mutant genome created by CRISPR-Cas9 and piggyBac recombinant technologies. Scientific Reports *8*.

[832] Liu, Y., Li, X., He, S., Huang, S., Li, C., Chen, Y., Liu, Z., Huang, X., and Wang, X. (2020). Efficient generation of mouse models with the prime editing system. Cell Discovery *6*.

[833] Lujan, S.A., Clausen, A.R., Clark, A.B., MacAlpine, H.K., MacAlpine, D.M., Malc, E.P., Mieczkowski, P.A., Burkholder, A.B., Fargo, D.C., Gordenin, D.A., *et al.* (2014). Heterogeneous polymerase fidelity and mismatch repair bias genome variation and composition. Genome Res *24*, 1751-1764.

[834] Mead, S., Whitfield, J., Poulter, M., Shah, P., Uphill, J., Campbell, T., Al-Dujaily, H., Hummerich, H., Beck, J., Mein, C.A., *et al.* (2009). A Novel Protective Prion Protein Variant that Colocalizes with Kuru Exposure. New England Journal of Medicine *361*, 2056-2065.

[835] Nelson, J.W., Randolph, P.B., Shen, S.P., Everette, K.A., Chen, P.J., Anzalone, A.V., Newby, G.A., An, M., Chen, J.C., and Liu, D.R. (2021). Engineered pegRNAs that improve prime editing efficiency. Submitted.

[836] Olson, H.E., Demarest, S.T., Pestana-Knight, E.M., Swanson, L.C., Iqbal, S., Lal, D., Leonard, H., Cross, J.H., Devinsky, O., and Benke, T.A. (2019). Cyclin-Dependent Kinase-Like 5 Deficiency Disorder: Clinical Review. Pediatric Neurology *97*, 18-25.

[837] Parsons, R., Li, G.-M., Longley, M.J., Fang, W.-H., Papadopoulos, N., Jen, J., De La Chapelle, A., Kinzler, K.W., Vogelstein, B., and Modrich, P. (1993). Hypermutability and mismatch repair deficiency in RER+ tumor cells. Cell *75*, 1227-1236.

[838] Petri, K., Zhang, W., Ma, J., Schmidts, A., Lee, H., Horng, J.E., Kim, D.Y., Kurt, I.C., Clement, K., Hsu, J.Y., *et al.* (2021). CRISPR prime editing with ribonucleoprotein complexes in zebrafish and primary human cells. Nature Biotechnology.

[839] Plotz, G., Raedle, J., Brieger, A., Trojan, J., and Zeuzem, S. (2003). N-terminus of hMLH1 confers interaction of hMutL and hMutL with hMutS. Nucleic Acids Research *31*, 3217-3226.

[840] Pluciennik, A., Dzantiev, L., Iyer, R.R., Constantin, N., Kadyrov, F.A., and Modrich, P. (2010). PCNA function in the activation and strand direction of MutL endonuclease in mismatch repair. Proceedings of the National Academy of Sciences *107*, 16066-16071.

[841] Ran, F.A., Cong, L., Yan, W.X., Scott, D.A., Gootenberg, J.S., Kriz, A.J., Zetsche, B., Shalem, O., Wu, X., Makarova, K.S., *et al.* (2015). In vivo genome editing using Staphylococcus aureus Cas9. Nature *520*, 186-191.

[842] Räschle, M., Dufner, P., Marra, G., and Jiricny, J. (2002). Mutations within the hMLH1 and hPMS2 Subunits of the Human MutLα Mismatch Repair Factor Affect Its ATPase Activity, but Not Its Ability to Interact with hMutSα. Journal of Biological Chemistry *277*, 21810-21820.

[843] Schene, I.F., Joore, I.P., Oka, R., Mokry, M., Van Vugt, A.H.M., Van Boxtel, R., Van Der Doef, H.P.J., Van Der Laan, L.J.W., Verstegen, M.M.A., Van Hasselt, P.M., *et al.* (2020). Prime editing for functional repair in patient-derived disease models. Nature Communications *11*.

[844] Shcherbakova, P.V., and Kunkel, T.A. (1999). Mutator phenotypes conferred by MLH1 overexpression and by heterozygosity for mlh1 mutations. Molecular and cellular biology *19*, 3177-3183.

[845] Sockolosky, J.T., Trotta, E., Parisi, G., Picton, L., Su, L.L., Le, A.C., Chhabra, A., Silveria, S.L., George, B.M., King, I.C., *et al.* (2018). Selective targeting of engineered T cells using orthogonal IL-2 cytokine-receptor complexes. Science *359*, 1037-1042.

[846] Spencer, J.M., and Zhang, X. (2017). Deep mutational scanning of S. pyogenes Cas9 reveals important functional domains. Scientific Reports *7*.

[847] Strand, M., Prolla, T.A., Liskay, R.M., and Petes, T.D. (1993). Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. Nature *365*, 274-276.

[848] Su, S.S., Lahue, R.S., Au, K.G., and Modrich, P. (1988). Mispair specificity of methyl-directed DNA mismatch correction in vitro. Journal of Biological Chemistry *263*, 6829-6835.

[849] Sugawara, N., Goldfarb, T., Studamire, B., Alani, E., and Haber, J.E. (2004). Heteroduplex rejection during single-strand annealing requires Sgs1 helicase and mismatch repair proteins Msh2 and Msh6 but not Pms1. Proceedings of the National Academy of Sciences *101*, 9315-9320.

[850] Supek, F., and Lehner, B. (2015). Differential DNA mismatch repair underlies mutation rate variation across the human genome. Nature *521*, 81-84.

[851] Sürün, D., Schneider, A., Mircetic, J., Neumann, K., Lansing, F., Paszkowski-Rogacz, M., Hänchen, V., Lee-Kirsch, M.A., and Buchholz, F. (2020). Efficient Generation and Correction of Mutations in Human iPS Cells Utilizing mRNAs of CRISPR Base Editors and Prime Editors. Genes *11*, 511.

[852] Thomas, D.C., Roberts, J.D., and Kunkel, T.A. (1991). Heteroduplex repair in extracts of human HeLa cells. Journal of Biological Chemistry *266*, 3744-3751.

[853] Tomer, G., Buermeyer, A.B., Nguyen, M.M., and Liskay, R.M. (2002). Contribution of human mlh1 and pms2 ATPase activities to DNA mismatch repair. J Biol Chem *277*, 21801-21809.

[854] Tran, H.T., Keen, J.D., Kricker, M., Resnick, M.A., and Gordenin, D.A. (1997). Hypermutability of homonucleotide runs in mismatch repair and DNA polymerase proofreading yeast mutants. Molecular and Cellular Biology *17*, 2859-2865.

[855] Trojan, J., Zeuzem, S., Randolph, A., Hemmerle, C., Brieger, A., Raedle, J., Plotz, G., Jiricny, J., and Marra, G. (2002). Functional analysis of hMLH1 variants and HNPCC-related mutations using a human expression system. Gastroenterology *122*, 211-219.

[856] Tsai, S.Q., Nguyen, N.T., Malagon-Lopez, J., Topkar, V.V., Aryee, M.J., and Joung, J.K. (2017). CIRCLE-seq: a highly sensitive in vitro screen for genome-wide CRISPR–Cas9 nuclease off-targets. Nature Methods *14*, 607-614.

[857] Umar, A., Boland, C.R., Terdiman, J.P., Syngal, S., Chapelle, A.D.L., Ruschoff, J., Fishel, R., Lindor, N.M., Burgart, L.J., Hamelin, R., *et al.* (2004). Revised Bethesda Guidelines

for Hereditary Nonpolyposis Colorectal Cancer (Lynch Syndrome) and Microsatellite Instability. JNCI Journal of the National Cancer Institute *96*, 261-268.

[858] Umar, A., Boyer, J.C., and Kunkel, T.A. (1994). DNA loop repair by human cell extracts. Science *266*, 814.

[859] Warren, J.J., Pohlhaus, T.J., Changela, A., Iyer, R.R., Modrich, P.L., and Beese, Lorena S. (2007). Structure of the Human MutSα DNA Lesion Recognition Complex. Molecular Cell *26*, 579-592.

[860] Wu, J., Corbett, A.H., and Berland, K.M. (2009). The Intracellular Mobility of Nuclear Import Receptors and NLS Cargoes. Biophysical Journal *96*, 3840-3849.

[861] Zhang, Y., Yuan, F., Presnell, S.R., Tian, K., Gao, Y., Tomkinson, A.E., Gu, L., and Li, G.-M. (2005). Reconstitution of 5′-Directed Human Mismatch Repair in a Purified System. Cell *122*, 693-705.

[862] Zhou, B.P., Liao, Y., Xia, W., Spohn, B., Lee, M.-H., and Hung, M.-C. (2001). Cytoplasmic localization of p21Cip1/WAF1 by Akt-induced phosphorylation in HER-2/neu-overexpressing cells. Nature Cell Biology *3*, 245-252.

EQUIVALENTS AND SCOPE

[863] In the articles such as "a," "an," and "the" may mean one or more than one unless indicated to the contrary or otherwise evident from the context. Embodiments or descriptions that include "or" between one or more members of a group are considered satisfied if one, more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process unless indicated to the contrary or otherwise evident from the context. The invention includes embodiments in which exactly one member of the group is present in, employed in, or otherwise relevant to a given product or process. The invention includes embodiments in which more than one, or all of the group members are present in, employed in, or otherwise relevant to a given product or process.

[864] Furthermore, the disclosure encompasses all variations, combinations, and permutations in which one or more limitations, elements, clauses, and descriptive terms from one or more of the listed claims is introduced into another claim. For example, any claim that is dependent on another claim can be modified to include one or more limitations found in any other claims that

is dependent on the same base claim. Where elements are presented as lists, *e.g.*, in Markush group format, each subgroup of the elements is also disclosed, and any element(s) can be removed from the group. It should it be understood that, in general, where the invention, or aspects of the invention, is/are referred to as comprising particular elements and/or features, certain embodiments of the disclosure or aspects of the disclosure consist, or consist essentially of, such elements and/or features. For purposes of simplicity, those embodiments have not been specifically set forth *in haec verba* herein. It is also noted that the terms "comprising" and "containing" are intended to be open and permits the inclusion of additional elements or steps. Where ranges are given, endpoints are included. Furthermore, unless otherwise indicated or otherwise evident from the context and understanding of one of ordinary skill in the art, values that are expressed as ranges can assume any specific value or sub–range within the stated ranges in different embodiments of the invention, to the tenth of the unit of the lower limit of the range, unless the context clearly dictates otherwise.

[865] This application refers to various issued patents, published patent applications, journal articles, and other publications, all of which are incorporated herein by reference. If there is a conflict between any of the incorporated references and the instant specification, the specification shall control. In addition, any particular embodiment of the present invention that falls within the prior art may be explicitly excluded from any one or more of the embodiments. Because such embodiments are deemed to be known to one of ordinary skill in the art, they may be excluded even if the exclusion is not set forth explicitly herein. Any particular embodiment of the invention can be excluded from any embodiment, for any reason, whether or not related to the existence of prior art.

[866] Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation many equivalents to the specific embodiments described herein. The scope of the present embodiments described herein is not intended to be limited to the above Description, but rather is as set forth in the appended embodiments. Those of ordinary skill in the art will appreciate that various changes and modifications to this description may be made without departing from the spirit or scope of the present invention, as defined in the following embodiments.

CLAIMS

What is claimed is:

1. A method for editing a nucleic acid molecule by prime editing comprising: contacting a nucleic acid molecule with a prime editor, a pegRNA, and an inhibitor of the DNA mismatch repair pathway, thereby installing one or more modifications to the nucleic acid molecule at a target site.

2. A method for editing a double stranded target DNA sequence, comprising: contacting the double stranded target DNA with (i) a prime editor, (ii) a prime editing guide RNA (pegRNA), and (iii) an inhibitor of a DNA mismatch repair pathway,

wherein the prime editor comprises a nucleic acid programmable DNA binding protein (napDNAbp) and a DNA polymerase,

wherein the pegRNA comprises a spacer sequence, a gRNA core, and an extension arm comprising a DNA synthesis template and a primer binding site (PBS),

wherein the spacer sequence comprises a region of complementarity to a target strand of the double stranded target DNA sequence,

wherein the gRNA core associates with the napDNAbp,

wherein the DNA synthesis template comprises a region of complementarity to the non-target strand of the double-stranded target DNA sequence and one or more nucleotide edits compared to the target strand double-stranded target DNA sequence;

wherein the primer binding site comprises a region of complementarity to a non-target strand of the double-stranded target DNA sequence,

wherein the contacting installs the one or more nucleotide edits in the double stranded target DNA, thereby editing the double stranded target DNA.

3. The method of claim 2, wherein the PBS comprises a region of complementarity to a region upstream of a nick site in the non-target strand of the target DNA sequence, wherein the nick site is characteristic of the napDNAbp.

4. The method of claim 1, wherein the method further comprises contacting the nucleic acid molecule with a second strand nicking gRNA.

5. The method of claim 1, wherein the prime editing efficiency is increased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of the DNA mismatch repair pathway compared to prime editing efficiency with the prime editor and the pegRNA in the absence of the inhibitor of the DNA mismatch repair pathway.

6. The method of claim 1, wherein the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-

fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of the DNA mismatch repair pathway compared to the frequency of indel formation with the prime editor and the pegRNA in the absence of the inhibitor of the DNA mismatch repair pathway.

7. 7. The method of claim 1, wherein the purity of editing outcome is increased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of the DNA mismatch repair pathway compared to the purity of editing outcome with the prime editor and the pegRNA in the absence of the inhibitor of the DNA mismatch repair pathway, wherein the purity of editing outcome is measured by the ratio of intended edit/unintended indels.

8. The method of claim 1, wherein the inhibitor of the DNA mismatch repair pathway inhibits the expression or function of one or more proteins of the DNA mismatch repair pathway (MMR proteins).

9. The method of claim 8, wherein the one or more MMR proteins is selected from the group consisting of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

10. The method of claim 8, wherein the one or more MMR proteins is MLH1.

11. The method of claim 10, wherein MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

12. The method of claim 8, wherein the inhibitor is an antibody that inhibits the activity of one or more proteins of the DNA mismatch repair pathway.

13. The method of claim 8, wherein the inhibitor is a small molecule that inhibits the activity of the one or more proteins of the DNA mismatch repair pathway.

14. The method of claim 8, wherein the inhibitor is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of the one or more proteins of the DNA mismatch repair pathway.

15. The method of claim 8, wherein the inhibitor is a dominant negative variant of an MMR protein that inhibits the activity of a wild type MMR protein.

16. The method of claim 8, wherein the inhibitor is a dominant negative variant of MLH1 that inhibits the activity of MLH1.

17. The method of claim 16, wherein the dominant negative variant of MLH1 comprises one or more amino acid substitutions, insertions, and/or deletions in an ATPase domain compared to

a wild type MLH1 protein as set forth in SEQ ID NO: 204, wherein the one or more amino acid alterations impairs or abolishes ATPase activity of the dominant negative variant of MLH1.

18. The method of claim 16, wherein the dominant negative variant of MLH1 comprises one or more amino acid substitutions, insertions, and/or deletions in an endonuclease domain compared to a wild type MLH1 protein as set forth in SEQ ID NO: 204, wherein the one or more amino acid alterations impairs or abolishes endonuclease activity of the dominant negative variant of MLH1.

19. The method of claim 16, wherein the dominant negative variant of MLH1 is truncated at the C terminus compared to a wild type MLH1 protein as set forth in SEQ ID NO: 204.

20. The method of claim 16, wherein the dominant negative variant of MLH1 is truncated at the N terminus compared to a wild type MLH1 protein as set forth in SEQ ID NO: 204.

21. The method of any one of claims 16-20, wherein the dominant negative variant of MLH1 further comprises a nuclear localization signal (NLS) at the N terminus and/or a NLS at the C terminus.

22. The method of claim 16, wherein the dominant negative variant is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.

23. The method of claim 16, wherein the dominant negative variant (a) comprises a E34A amino acid substitution; (b) a deletion of amino acid 756; (c) a deletion of amino acids 754-756; (d)

a E34A amino acid substitution and a deletion of amino acids 754-756; (e) a deletion of amino acids 336-756; (f) a E34A amino acid substitution and a deletion of amino acids 336-756; (g) a deletion of amino acids 1-500; (h) a deletion of amino acids 1-500 and a deletion of amino acids 754-756; or (i) a deletion of amino acids 1-460 and a deletion of amino acids 754-756, optionally wherein the dominant negative variant further comprises an NLS comprising the sequence KRTADGSEFESPKKKRKV at the C terminus and/or at the N terminus.

24. The method of claim 8, wherein the inhibitor is a dominant negative variant of PMS2 that inhibits the activity of PMS2, optionally wherein the dominant negative variant comprises one or more amino acid substitutions, insertions, and/or deletions compared to a wild type PMS2 protein, optionally wherein the dominant negative variant comprises (a) a E705K substitution, (b) a deletion of amino acids 2-607, (c) a deletion of amino acids 2-635, (d) a deletion of amino acids 1-635, (e) a E41A substitution, and/or (f) a deletion after amino acid 134 compared to a wild type PMS2 protein.

25. The method of claim 8, wherein the inhibitor is a dominant negative variant of MSH6 that inhibits the activity of MSH6, optionally wherein the dominant negative variant comprises one or more amino acid substitutions, insertions, and/or deletions compared to a wild type MSH6 protein, optionally wherein the dominant negative variant comprises (a) a K1140R substitution, and/or (b) a deletion of amino acids 2-361 compared to a wild type MSH6 protein.

26. The method of claim 8, wherein the inhibitor comprises CDKN1A.

27. The method of claim 1, wherein the prime editor comprises a napDNAbp and a polymerase.

28. The method of claim 27, wherein the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof.

29. The method of claim 27, wherein the napDNAbp is a Cas9 nickase comprising one or more amino acid substitutions in the HNH domain, optionally wherein the one or more amino acid substitutions comprises H840X, N854X, and/or N863X, wherein X is any amino acid except

for the original amino acid, optionally wherein the one or more amino acid substitutions comprises H840A, N854A, and/or N863A.

30. The method of claim 27, wherein the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, Cas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute and optionally has a nickase activity.

31. The method of claim 27, wherein the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, or 104, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2,4-67, or 104.

32. The method of claim 27, wherein the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 37 (i.e., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2 or SEQ ID NO: 37.

33. The method of claim 27, wherein the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase.

34. The method of claim 27, wherein the polymerase is a reverse transcriptase.

35. The method of claim 34, wherein the reverse transcriptase is a retroviral reverse transcriptase, optionally wherein the reverse transcriptase is a Moloney Murine Leukemia virus reverse transcriptase (MMLV-RT), optionally wherein the MMLV-RT comprises one or more amino acid substitutions selected from D200N, T306K, W313F, T330P, and L603W compared to a wild type MMLV-RT.

36. The method of claim 34, wherein the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98, optionally wherein the reverse transcriptase comprises an amino acid sequence of SEQ ID NO: 105.

37. The method of claim 27, wherein the napDNAbp and the polymerase of the prime editor are joined to form a fusion protein, optionally wherein the napDNAbp and the polymerase are joined by a linker.

38. The method of claim 37, wherein the linker comprises an amino acid sequence of any one of SEQ ID NOs: 102 or 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102 or 118-131.

39. The method of claim 37, wherein the linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

40. The method of claim 37, wherein the fusion protein comprises the amino acid sequence of SEQ ID NO: 99 or SEQ ID NO: 107, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NO: 99 or SEQ ID NO: 107.

41. The method of claim 1, wherein the prime editor, the pegRNA, and the inhibitor of the DNA mismatch repair pathway are encoded on one or more DNA vectors.

42. The method of claim 41, wherein the one or more DNA vectors comprise AAV or lentivirus DNA vectors.

43. The method of claim 42, wherein the AAV vector is serotype 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

44. The method of any one of claims 27-36, wherein the prime editor and the inhibitor of the DNA mismatch repair pathway are not covalently linked.

45. The method of any one of claims 27-36, wherein the napDNAbp, the polymerase, or the prime editor as a fusion protein is further joined by a second linker to the inhibitor of the DNA mismatch repair pathway.

46. The method of claim 45, wherein the second linker comprises a self-hydrolyzing linker, optionally wherein the second linker is a T2A linker or a P2A linker.

47. The method of claim 45, wherein the second linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or 233-236, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131, or 233-236.

48. The method of claim 45, wherein the second linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

49. The method of claim 1, wherein the one or more modifications to the nucleic acid molecule installed at the target site comprises one or more transitions, one or more transversions, one or more insertions, one or more deletions, one more inversions, or any combination thereof, and optionally are less than 15 bp.

50. The method of claim 49, wherein the one or more transitions are selected from the group consisting of: (a) T to C; (b) A to G; (c) C to T; and (d) G to A.

51. The method of claim 49, wherein the one or more transversions are selected from the group consisting of: (a) T to A; (b) T to G; (c) C to G; (d) C to A; (e) A to T; (f) A to C; (g) G to C; and (h) G to T.

52. The method of claim 1, wherein the one or more modifications comprises changing (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair to a C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair.

53. The method of claim 1, wherein the one or more modifications comprises an insertion or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides, optionally wherein the one or more edits comprises an insertion or deletion of 1-15 nucleotides.

54. The method claim 1, wherein the one or more modifications comprises a correction to a mutation associated with a disease in a disease-associated gene.

55. The method of claim 54, wherein the disease-associated gene is associated with a polygenic disorder selected from the group consisting of: heart disease; high blood pressure; Alzheimer's disease; arthritis; diabetes; cancer; and obesity.

56. The method of claim 54, wherein the disease-associated gene is associated with a monogenic disorder selected from the group consisting of: Adenosine Deaminase (ADA) Deficiency; Alpha-1 Antitrypsin Deficiency; Cystic Fibrosis; Duchenne Muscular Dystrophy; Galactosemia; Hemochromatosis; Huntington's Disease; Maple Syrup Urine Disease; Marfan Syndrome; Neurofibromatosis Type 1; Pachyonychia Congenita; Phenylketonuria; Severe Combined Immunodeficiency; Sickle Cell Disease; Smith-Lemli-Opitz Syndrome; a trinucleotide repeat disorder; a prion disease; and Tay-Sachs Disease.

57. The method of any one of claims 1-56, wherein the gRNA core comprises minimal sequence homology to the sequence of the target site, optionally wherein the gRNA core comprises no more than 1%, 5%, 10%, 15%, 20%, 25%, or 30% sequence homology to the sequence of the double stranded target DNA that flanks 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 nucleotides upstream or downstream of the position of the one or more nucleotide edits.

58. A composition for editing a nucleic acid molecule by prime editing comprising a prime editor, a pegRNA, and an inhibitor of the DNA mismatch repair pathway, wherein the composition is capable of installing one or more modifications to the nucleic acid molecule at a target site.

59. The composition of claim 58, wherein the composition further comprises a second strand nicking gRNA.

60. The composition of claim 58, wherein the prime editing efficiency is increased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold,

at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold in the presence of the inhibitor of the DNA mismatch repair pathway compared to prime editing efficiency with the prime editor and the pegRNA in the absence of the inhibitor of the DNA mismatch repair pathway.

61. The composition of claim 58, wherein the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at

least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold in the presence of the inhibitor of the DNA mismatch repair pathway compared to the frequency of indel formation with the prime editor and the pegRNA in the absence of the inhibitor of the DNA mismatch repair pathway.

62. The composition of claim 58, wherein the purity of editing outcome is increased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of the DNA mismatch repair pathway compared to the purity of editing outcome with the prime editor and the pegRNA in the absence of the inhibitor of the DNA mismatch repair pathway, wherein the purity of editing outcome is measured by the ratio of intended edit/unintended indels.

63. The composition of claim 58, wherein the inhibitor of the DNA mismatch repair pathway inhibits the expression of one or more proteins of the DNA mismatch repair pathway (MMR proteins).

64. The composition of claim 63, wherein the one or more MMR proteins is selected from the group consisting of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL

gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

65. The composition of claim 63, wherein the one or more MMR proteins is MLH1.

66. The composition of claim 65, wherein MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

67. The composition of claim 63, wherein the inhibitor is an antibody that inhibits the activity of one or more proteins of the DNA mismatch repair pathway.

68. The composition of claim 63, wherein the inhibitor is a small molecule that inhibits the activity of the one or more proteins of the DNA mismatch repair pathway.

69. The composition of claim 63, wherein the inhibitor is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of the one or more proteins of the DNA mismatch repair pathway.

70. The composition of claim 63, wherein the inhibitor is a dominant negative variant of an MMR protein that inhibits the activity of a wild type MMR protein.

71. The composition of claim 63, wherein the inhibitor is a dominant negative variant of MLH1 that inhibits the activity of MLH1.

72. The composition of claim 71, wherein the dominant negative variant of MLH1 comprises one or more amino acid substitutions, insertions, and/or deletions in an ATPase domain compared to a wild type MLH1 protein as set forth in SEQ ID NO: 204, wherein the one or more amino acid alterations impairs or abolishes ATPase activity of the dominant negative variant of MLH1.

73. The composition of claim 71, wherein the dominant negative variant of MLH1 comprises one or more amino acid substitutions, insertions, and/or deletions in an endonuclease domain

compared to a wild type MLH1 protein as set forth in SEQ ID NO: 204, wherein the one or more amino acid alterations impairs or abolishes endonuclease activity of the dominant negative variant of MLH1.

74. The composition of claim 71, wherein the dominant negative variant of MLH1 is truncated at the C terminus compared to a wild type MLH1 protein as set forth in SEQ ID NO: 204.

75. The composition of claim 71, wherein the dominant negative variant of MLH1 is truncated at the N terminus compared to a wild type MLH1 protein as set forth in SEQ ID NO: 204.

76. The composition of any one of claims 71-75, wherein the dominant negative variant of MLH1 further comprises a nuclear localization signal (NLS) at the N terminus and/or a NLS at the C terminus.

77. The composition of claim 71, wherein the dominant negative variant is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.

78. The composition of claim 71, wherein the dominant negative variant (a) comprises a E34A amino acid substitution; (b) a deletion of amino acid 756; (c) a deletion of amino acids 754-756; (d) a E34A amino acid substitution and a deletion of amino acids 754-756; (e) a deletion of amino acids 336-756; (f) a E34A amino acid substitution and a deletion of amino acids 336-756; (g) a deletion of amino acids 1-500; (h) a deletion of amino acids 1-500 and a deletion of amino acids 754-756; or (i) a deletion of amino acids 1-460 and a deletion of amino acids 754-756, optionally wherein the dominant negative variant further comprises an

NLS comprising the sequence KRTADGSEFESPKKKRKV at the C terminus and/or at the N terminus.

79. The composition of claim 73, wherein the inhibitor is a dominant negative variant of PMS2 that inhibits the activity of PMS2, optionally wherein the dominant negative variant comprises one or more amino acid substitutions, insertions, and/or deletions compared to a wild type PMS2 protein, optionally wherein the dominant negative variant comprises (a) a E705K substitution, (b) a deletion of amino acids 2-607, (c) a deletion of amino acids 2-635, (d) a deletion of amino acids 1-635, (e) a E41A substitution, and/or (f) a deletion after amino acid 134 compared to a wild type PMS2 protein.

80. The composition of claim 73, wherein the inhibitor is a dominant negative variant of MSH6 that inhibits the activity of MSH6, optionally wherein the dominant negative variant comprises one or more amino acid substitutions, insertions, and/or deletions compared to a wild type MSH6 protein, optionally wherein the dominant negative variant comprises (a) a K1140R substitution, and/or (b) a deletion of amino acids 2-361 compared to a wild type MSH6 protein.

81. The composition of claim 73, wherein the inhibitor comprises CDKN1A.

82. The composition of claim 58, wherein the prime editor comprises a napDNAbp and a polymerase.

83. The composition of claim 82, wherein the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof.

84. The composition of claim 82, wherein the napDNAbp is a Cas9 nickase comprising one or more amino acid substitutions in the HNH domain, optionally wherein the one or more amino acid substitutions comprises H840X, N854X, and/or N863X, wherein X is any amino acid except for the original amino acid, optionally wherein the one or more amino acid substitutions comprises H840A, N854A, and/or N863A.

85. The composition of claim 84, wherein the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, ArgonauteCas12b2, Cas13a,

Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute and optionally has a nickase activity.

86. The composition of claim 84, wherein the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, or PEmax or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2, 4-67, or 104.

87. The composition of claim 84, wherein the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 37 (i.e., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2 or SEQ ID NO: 37.

88. The composition of claim 84, wherein the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase.

89. The composition of claim 84, wherein the polymerase is a reverse transcriptase.

90. The composition of claim 89, wherein the reverse transcriptase is a retroviral reverse transcriptase, optionally wherein the reverse transcriptase is a Moloney Murine Leukemia virus reverse transcriptase (MMLV-RT), optionally wherein the MMLV-RT comprises one or more amino acid substitutions selected from D200N, T306K, W313F, T330P, and L603W compared to a wild type MMLV-RT.

91. The composition of claim 89, wherein the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98, optionally wherein the reverse transcriptase comprises an amino acid sequence of SEQ ID NO: 105.

92. The composition of claim 84, wherein the napDNAbp and the polymerase of the prime editor are joined to form a fusion protein, optionally wherein the napDNAbp and the polymerase are joined by a linker.

93. The composition of claim 92, wherein the linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131.

94. The composition of claim 92, wherein the linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

95. The composition of claim 92, wherein the fusion protein comprises the amino acid sequence of SEQ ID NO: 99 or SEQ ID NO: 107, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NO: 99 or SEQ ID NO: 107.

96. The composition of claim 58, wherein the prime editor, the pegRNA, and the inhibitor of the DNA mismatch repair pathway are encoded on one or more DNA vectors.

97. The composition of claim 96, wherein the one or more DNA vectors comprise AAV or lentivirus DNA vectors.

98. The composition of claim 97, wherein the AAV vector is serotype 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

99. The composition of any one of claims 84-91, wherein the prime editor and the inhibitor of the DNA mismatch repair pathway are not covalently linked.

100. The composition of any one of claims 84-91, wherein the prime editor as a fusion protein is further joined by a second linker to the inhibitor of the DNA mismatch repair pathway.

101. The composition of claim 100, wherein the second linker comprises a self-hydrolyzing linker, optionally wherein the second linker is a T2A linker or a P2A linker.

102. The composition of claim 100, wherein the second linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or 233-236, or an amino acid sequence

having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131, 233-236.

103. The composition of claim 100, wherein the second linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

104. The composition of claim 58, wherein the one or more modifications to the nucleic acid molecule installed at the target site comprises one or more transitions, one or more transversions, one or more insertions, one or more deletions, one more inversions, or any combination thereof, and optionally are less than 15 bp.

105. The composition of claim 104, wherein the one or more transitions are selected from the group consisting of: (a) T to C; (b) A to G; (c) C to T; and (d) G to A.

106. The composition of claim 104, wherein the one or more transversions are selected from the group consisting of: (a) T to A; (b) T to G; (c) C to G; (d) C to A; (e) A to T; (f) A to C; (g) G to C; and (h) G to T.

107. The composition of claim 58, wherein the one or more modifications comprises changing (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair to a C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair.

108. The composition of claim 58, wherein the one or more modifications comprises an insertion or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides, optionally wherein the one or more edits comprises an insertion or deletion of 1-15 nucleotides.

109. The composition claim 58, wherein the one or more modifications comprises a correction to a mutation associated with a disease in a disease-associated gene.

110. The composition of claim 109, wherein the disease-associated gene is associated with a polygenic disorder selected from the group consisting of: heart disease; high blood pressure; Alzheimer's disease; arthritis; diabetes, cancer; and obesity.

111. The composition of claim 109, wherein the disease-associated gene is associated with a monogenic disorder selected from the group consisting of: Adenosine Deaminase (ADA) Deficiency; Alpha-1 Antitrypsin Deficiency; Cystic Fibrosis; Duchenne Muscular Dystrophy; Galactosemia; Hemochromatosis; Huntington's Disease; Maple Syrup Urine Disease; Marfan Syndrome; Neurofibromatosis Type 1; Pachyonychia Congenita; Phenylketonuria; Severe Combined Immunodeficiency; Sickle Cell Disease; Smith-Lemli-Opitz Syndrome; a trinucleotide repeat disorder; a prion disease; and Tay-Sachs Disease.

112. The composition of any one of claims 58-111, wherein the gRNA core comprises minimal sequence homology to the sequence of the target site, optionally wherein the gRNA core comprises no more than 1%, 5%, 10%, 15%, 20%, 25%, or 30% sequence homology to the sequence of the double stranded target DNA that flanks 5, 10, 15, 20, 25, 30, 35, 40, 45, or 50 nucleotides upstream or downstream of the position of the one or more nucleotide edits.

113. A polynucleotide for editing a DNA target site by prime editing comprising a nucleic acid sequence encoding a napDNAbp, a polymerase, and an inhibitor of the DNA mismatch repair pathway, wherein the napDNAbp and polymerase is capable in the presence of a pegRNA of installing one or more modifications in the DNA target site.

114. The polynucleotide of claim 113, wherein the polynucleotide further comprises a nucleic acid sequence encoding a second strand nicking gRNA.

115. The polynucleotide of claim 113, wherein the prime editing efficiency is increased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at

least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of the DNA mismatch repair pathway.

116. The polynucleotide of claim 113, wherein the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of the DNA mismatch repair pathway.

117. The polynucleotide of claim 113, wherein the inhibitor of the DNA mismatch repair pathway inhibits one or more proteins of the DNA mismatch repair pathway.

118. The polynucleotide of claim 117, wherein the one or more proteins is selected from the group consisting of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

119. The polynucleotide of claim 117, wherein the one or more proteins is MLH1.

120. The polynucleotide of claim 119, wherein MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

121. The polynucleotide of claim 113, wherein the inhibitor is an antibody that inhibits the activity of one or more proteins of the DNA mismatch repair pathway.

122. The polynucleotide of claim 113, wherein the inhibitor is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of one or more proteins of the DNA mismatch repair pathway.

123. The polynucleotide of claim 113, wherein the inhibitor is a dominant negative variant of an MMR protein that inhibits the activity of a wild type MMR protein.

124. The polynucleotide of claim 113, wherein the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

125. The polynucleotide of claim 124, wherein the dominant negative variant is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least

99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.

126. The polynucleotide of claim 113, wherein the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof.

127. The polynucleotide of claim 113, wherein the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, ArgonauteCas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute and optionally has a nickase activity.

128. The polynucleotide of claim 113, wherein the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, or 104 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2,4-67, or 104.

129. The polynucleotide of claim 113, wherein the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 (i.e., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2.

130. The polynucleotide of claim 113, wherein the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase.

131. The polynucleotide of claim 113, wherein the polymerase is a reverse transcriptase.

132. The polynucleotide of claim 131, wherein the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98.

133. The polynucleotide of claim 113, wherein the napDNAbp and the polymerase of the prime editor are joined by a linker to form a fusion protein.

134. The polynucleotide of claim 133, wherein the linker comprises an amino acid sequence of any one of SEQ ID NOs: 102 or 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102 or 118-131.

135. The polynucleotide of claim 133, wherein the linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

136. The polynucleotide of claim 113, wherein the polynucleotide is a DNA vector.

137. The polynucleotide of claim 136, wherein the DNA vector is an AAV or lentivirus DNA vector.

138. The polynucleotide of claim 137, wherein the AAV vector is serotype 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

139. The polynucleotide of claim 133, wherein the prime editor as a fusion protein is further joined by a second linker to the inhibitor of the DNA mismatch repair pathway.

140. The polynucleotide of claim 139, wherein the second linker comprises a self-hydrolyzing linker.

141. The polynucleotide of claim 139, wherein the second linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or 233-236, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131, or 233-236.

142. The polynucleotide of claim 139, wherein the second linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

143. The polynucleotide of claim 113, wherein the one or more modifications to the nucleic acid molecule installed at the target site comprises one or more transitions, one or more transversions, one or more insertions, one or more deletions, or one more inversions.

144. The polynucleotide of claim 143, wherein the one or more transitions are selected from the group consisting of: (a) T to C; (b) A to G; (c) C to T; and (d) G to A.

145. The polynucleotide of claim 143, wherein the one or more transversions are selected from the group consisting of: (a) T to A; (b) T to G; (c) C to G; (d) C to A; (e) A to T; (f) A to C; (g) G to C; and (h) G to T.

146. A polynucleotide of claim 113, wherein the one or more modifications comprises changing (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair to a C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair.

147. A polynucleotide of claim 113, wherein the one or more modifications comprises an insertion or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides.

148. The polynucleotide of claim 113, wherein the one or more modifications comprises a correction to a disease-associated gene.

149. The polynucleotide of claim 148, wherein the disease-associated gene is associated with a polygenic disorder selected from the group consisting of: heart disease; high blood pressure; Alzheimer's disease; arthritis; diabetes; cancer; and obesity.

150. The polynucleotide of claim 148, wherein the disease-associated gene is associated with a monogenic disorder selected from the group consisting of: Adenosine Deaminase (ADA) Deficiency; Alpha-1 Antitrypsin Deficiency; Cystic Fibrosis; Duchenne Muscular Dystrophy; Galactosemia; Hemochromatosis; Huntington's Disease; Maple Syrup Urine Disease; Marfan Syndrome; Neurofibromatosis Type 1; Pachyonychia Congenita; Phenylketonuria; Severe Combined Immunodeficiency; Sickle Cell Disease; Smith-Lemli-Opitz Syndrome; a trinucleotide repeat disorder; a prion disease; and Tay-Sachs Disease.

151. A cell comprising a polynucleotide of any of claims 113-150, optionally wherein the cell is a mammalian cell, a non-human primate cell, or a human cell.

152. A pharmaceutical composition comprising the composition of any of claims 41-80 or the polynucleotide of any of claims 113-151, or the cell of claim 151, and a pharmaceutical excipient.

153. A kit comprising the composition of any of claims 58-112 or the polynucleotide of any of claims 113-150, a pharmaceutical excipient, and instructions for editing a DNA target site by prime editing.

154. A composition comprising a first nucleic acid sequence encoding a nucleic acid programmable DNA binding protein (napDNAbp), a second nucleic acid sequence encoding a polymerase, and a third nucleic acid sequence encoding an inhibitor of the DNA mismatch repair pathway.

155. The composition of claim 154, wherein the composition further comprises a prime editing guide RNA (pegRNA) or a nucleic acid sequence encoding the pegRNA, wherein the pegRNA comprises a spacer sequence, a gRNA core, and an extension arm comprising a DNA synthesis template and a primer binding site (PBS), wherein the spacer sequence comprises a region of complementarity to a target strand of a double stranded target DNA sequence, wherein the gRNA core associates with the napDNAbp, wherein the DNA synthesis template comprises a region of complementarity to the non-target strand of the double-stranded target DNA sequence and one or more nucleotide edits compared to the target strand double-stranded target DNA sequence, and wherein the primer binding site comprises a region of complementarity to a non-target strand of the double-stranded target DNA sequence.

156. The composition of claim 154, wherein the first and the second nucleic acid sequences are on a single polynucleotide.

157. The composition of claim 154, wherein the first, the second, and the third nucleic acid sequences are on a single polynucleotide.

158. The composition of claim 157, wherein the first and the second nucleic acid sequences are connected to encode a napDNAbp-DNA polymerase fusion protein.

159. The composition of claim 158, wherein the first and the second nucleic acid sequences are connected to encode a napDNAbp-DNA polymerase fusion protein, wherein the third nucleic acid sequence is connected to the first or the second nucleic acide sequence via a linker nucleic acid sequence, optionally wherein the linker nucleic acid sequence encodes a peptide linker, optionally wherein the peptide linker is a self-hydrolyzing linker, optionally wherein the self-hydrolyzing linker is a T2A linker or a P2A linker, optionally wherein the self-hydrolyzing linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or 233-236, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131, or 233-236.

160. The composition of claim 156 or 157, wherein the single polynucleotide is part of a DNA vector.

161. The composition of claim 156 or 157, wherein the single polynucleotide is part of an mRNA sequence.

162. The composition of claim 160, wherein the DNA vector is an AAV or lentivirus DNA vector, optionally wherein the DNA vector further comprises a promoter.

163. The composition of claim 161, wherein the mRNA sequence further comprises a promoter.

164. A method for editing a nucleic acid molecule by prime editing comprising: contacting a nucleic acid molecule with a prime editor and a prime editing guide RNA (pegRNA), wherein the prime editor comprises a nucleic acid programmable DNA binding protein and a DNA polymerase, wherein the pegRNA a spacer sequence, a gRNA core, and an extension arm comprising a DNA synthesis template and a primer binding site (PBS), wherein the DNA synthesis template comprises three or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule, wherein the three or more consecutive nucleotide mismatches comprise (i) an insertion, deletion, or substitution of x

consecutive nucleotides that corrects a mutation (e.g., a disease associated mutation) and (ii) an insertion, deletion, or substitution of y consecutive nucleotides directly adjacent to the x consecutive nucleotides, wherein the insertion, deletion, or substitution of y consecutive nucleotides is a silent mutation, wherein (x+y) is an integer no less than 3, wherein y is an integer no less than 1, and wherein inclusion of the silent mutation(s) increases the efficiency, reduces unintended indel frequency, and/or improves editing outcome purity by prime editing.

165. The method of claim 164, wherein at least one of the three or more consecutive nucleotide mismatches results in an alteration in the amino acid sequence of a protein expressed from the nucleic acid molecule, and wherein at least one of the remaining three or more consecutive nucleotide mismatches are silent mutations.

166. The method of claim 165, wherein the silent mutations are in a coding region of the target nucleic acid molecule.

167. The method of claim 166, wherein the silent mutations introduce into the nucleic acid molecule one or more alternate codons encoding the same amino acid as the unedited nucleic acid molecule.

168. The method of claim 165, wherein the silent mutations are in a non-coding region of the target nucleic acid molecule.

169. The method of claim 168, wherein the silent mutations do not influence splicing, gene regulation, RNA lifetime, or other biological properties of the target site on the nucleic acid molecule.

170. The method of any one of claims 164-169, wherein the DNA synthesis template of the pegRNA comprises four or more, five or more, six or more, seven or more, eight or more, nine or more, or ten or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule.

171. The method of any one of claims 164-170, wherein the three or more consecutive nucleotide mismatches evade correction by the DNA mismatch repair pathway.

172. The method of any one of claims 164-171, wherein the prime editing efficiency is increased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold relative to a method using a pegRNA comprising a DNA synthesis template comprising only one consecutive nucleotide mismatch relative to the endogenous sequence of a target site on the nucleic acid molecule.

173. The method of any one of claims 164-171, wherein the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-

fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold relative to a method using a pegRNA comprising a DNA synthesis template comprising only one consecutive nucleotide mismatch relative to the endogenous sequence of a target site on the nucleic acid molecule.

174. A method for editing a nucleic acid molecule by prime editing comprising: contacting a nucleic acid molecule with a prime editor and a pegRNA, wherein the extension arm of the pegRNA comprises a DNA synthesis template comprising an insertion or deletion of 10 or more nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule.

175. The method of claim 174, wherein the DNA synthesis template comprises an insertion or deletion of 11 or more nucleotides, 12 or more nucleotides, 13 or more nucleotides, 14 or more nucleotides, 15 or more nucleotides, 16 or more nucleotides, 17 or more nucleotides, 18 or more nucleotides, 19 or more nucleotides, 20 or more nucleotides, 21 or more nucleotides, 22 or more nucleotides, 23 or more nucleotides, 24 or more nucleotides, or 25 or more nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule.

176. The method of claim 175, wherein the DNA synthesis template comprises an insertion or deletion of 15 or more nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule.

177. The method of any one of claims 174-176, wherein the insertion or deletion of 10 or more nucleotides relative to the endogenous sequence of the target site on the nucleic acid molecule evades correction by the DNA mismatch repair pathway.

178. The method of any one of claims 174-176, wherein the prime editing efficiency is increased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a method using a pegRNA comprising a DNA synthesis

template comprising an insertion or deletion of fewer than 10 nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule.

179. The method of claim 174-176, wherein the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold relative to a method using a pegRNA comprising a DNA synthesis template comprising an insertion or deletion of fewer than 10 nucleotides relative to the endogenous sequence of a target site on the nucleic acid molecule.

180. A prime editing guide RNA (pegRNA) for editing a nucleic acid molecule by prime editing, wherein the pegRNA comprises a spacer sequence, a gRNA core, and an extension arm comprising a DNA synthesis template and a primer binding site (PBS), wherein the DNA synthesis template comprising three or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule, wherein the three or more consecutive nucleotide mismatches comprise (i) an insertion, deletion, or substitution of x nucleotides that corrects a mutation (e.g. a disease associated mutation) and (ii) an insertion, deletion, or substitution of y nucleotides directly adjacent to the x nucleotides, wherein the insertion, deletion, or substitution of y nucleotides is a silent mutation, wherein (x+y) is an integer no less than 3, wherein y is an integer no less than 1, and wherein inclusion of the silent

mutation(s) increases the efficiency, reduces unintended indel frequency, and/or improves editing outcome purity by prime editing.

181. The pegRNA of claim 180, wherein at least one of the three or more consecutive nucleotide mismatches results in an alteration in the amino acid sequence of a protein expressed from the nucleic acid molecule, and wherein at least one of the remaining three or more consecutive nucleotide mismatches are silent mutations.

182. The pegRNA of claim 181, wherein the silent mutations are in a coding region of the target nucleic acid molecule.

183. The pegRNA of claim 182, wherein the silent mutations introduce into the nucleic acid molecule one or more alternate codons encoding the same amino acid as the unedited nucleic acid molecule.

184. The pegRNA of claim 181, wherein the silent mutations are in a non-coding region of the target nucleic acid molecule.

185. The pegRNA of claim 183, wherein the silent mutations are in a region of the nucleic acid molecule that does not influence splicing, gene regulation, RNA lifetime, or other biological properties of the target site on the nucleic acid molecule.

186. The pegRNA of any one of claims 180-185, wherein the extension arm of the pegRNA comprises four or more, five or more, six or more, seven or more, eight or more, nine or more, or ten or more consecutive nucleotide mismatches relative to the endogenous sequence of a target site on the nucleic acid molecule.

187. The pegRNA of any one of claims 180-186, wherein the three or more consecutive nucleotide mismatches evade the DNA mismatch repair pathway.

188. The pegRNA of any one of claims 180-186, wherein use of the pegRNA in prime editing results in the prime editing efficiency being increased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least

8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold relative to a pegRNA comprising a DNA synthesis template comprising only one consecutive nucleotide mismatch relative to the endogenous sequence of a target site on the nucleic acid molecule.

189. The pegRNA of any one of claims 180-186, wherein use of the pegRNA in prime editing results in the frequency of indel formation being decreased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0-fold relative to a pegRNA comprising a DNA synthesis template comprising only one consecutive nucleotide mismatch relative to the endogenous sequence of a target site on the nucleic acid molecule.

190. A prime editor system comprising the pegRNA of any one of claims 180-189 and a prime editor, wherein the prime editor comprises a napDNAbp and a polymerase.

191. The prime editor system of claim 190, further comprising an inhibitor of DNA mismatch repair pathway.

192. A prime editor comprising a fusion protein comprising (i) a nucleic acid programmable DNA binding protein (napDNAbp) and (ii) a DNA polymerase, wherein the napDNAbp is a

Cas9 nickase (nCas9) comprising a R221K amino acid substitution, a N39K amino acid substitution, and an amino acid substitution that inactivates HNH domain nuclease activity, or corresponding amino acid substitutions thereof, relative to a wild type Cas9 as set forth in SEQ ID NO: 2.

193. The prime editor of claim 192, wherein the nCas9 comprises a R221K, a N39K, and a H840A amino acid substitution compared to a wild type Cas9 as set forth in SEQ ID NO: 2.

194. The prime editor of claim 193, wherein the nCas9 and the DNA polymerase are connected by a linker, optionally wherein the linker comprises the sequence of SEQ ID NO: X5, optionally wherein the prime editor further comprises a SV40 NLS at the N terminus, optionally wherein the prime editor further comprises a SV40 NLS and/or a c-Myc NLS at the C terminus.

195. The prime editor of claim 192 comprising the amino acid sequence of SEQ ID NO: 99 or an amino acid sequence have at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% sequence identity with SEQ ID NO: 99.

196. A prime editor system comprising the prime editor of any one of claims 192-195 and an inhibitor of DNA mismatch repair pathway.

197. A prime editor system comprising the prime editor of any one of claims 192-195 and a prime editing guide RNA (pegRNA).

198. A polynucleotide encoding the prime editor of any of claims 192-195.

199. The polynucleotide of claim 198, wherein the polynucleotide is DNA.

200. The polynucleotide of claim 198, wherein the polynucleotide is mRNA.

201. A vector comprising the polynucleotide of claim 198, optionally wherein expression of the fusion protein is under the control of a promoter, optionally wherein the promoter is a U6 promoter.

202. A prime editor system for site specific genome modification, comprising a (a) prime editor comprising (i) a nucleic acid programmable DNA binding protein (napDNAbp) and (ii) a DNA polymerase, and (b) an inhibitor of the DNA mismatch repair pathway.

203. The prime editor system of claim 202, wherein the inhibitor of the DNA mismatch repair pathway inhibits one or more proteins of the DNA mismatch repair pathway.

204. The prime editor system of claim 203, wherein the one or more proteins is selected from the group consisting of MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

205. The prime editor system of claim 203, wherein the one or more proteins is MLH1.

206. The prime editor system of claim 205, wherein the MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

207. The prime editor system of claim 202, wherein the inhibitor is an antibody that inhibits the activity of one or more proteins of the DNA mismatch repair pathway.

208. The prime editor system of claim 202, wherein the inhibitor is a small molecule that inhibits the activity of one or more proteins of the DNA mismatch repair pathway.

209. The prime editor system of claim 202, wherein the inhibitor is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of one or more proteins of the DNA mismatch repair pathway.

210. The prime editor system of claim 202, wherein the inhibitor is a dominant negative variant of an MMR protein that inhibits the activity of a wild type MMR protein.

211. The prime editor system of claim 202, wherein the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

212. The prime editor system of claim 202, wherein the dominant negative variant is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.

213. The prime editor system of claim 202, wherein the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof.

214. The prime editor system of claim 202, wherein the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, ArgonauteCas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute and optionally has a nickase activity.

215. The prime editor system of claim 202, wherein the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, 104 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2, 4-67, 104.

216. The prime editor system of claim 202, wherein the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 (i.e., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2.

217. The prime editor system of claim 202, wherein the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase.

218. The prime editor system of claim 202, wherein the polymerase is a reverse transcriptase.

219. The prime editor system of claim 218, wherein the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98.

220. The prime editor system of claim 202, wherein the napDNAbp and the polymerase of the prime editor are joined by a linker to form a fusion protein, and optionally wherein the inhibitor is joined by a second linker to either the napDNAbp or the polymerase.

221. The prime editor system of claim 220, wherein the linker and/or second linker comprises an amino acid sequence of any one of SEQ ID NOs: 102 or 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102 or 118-131.

222. The prime editor system of claim 220, wherein the linker and/or second linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

223. The prime editor system of claim 202, wherein the prime editor is PE1 of SEQ ID NO: 100 or an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or up to 100% sequence identity with SEQ ID NO: 100.

224. The prime editor system of claim 202, wherein the prime editor is PE2 of SEQ ID NO: 107 or an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or up to 100% sequence identity with SEQ ID NO: 107.

225. The prime editor system of claim 202, wherein the prime editor is PE1 of SEQ ID NO: 100 or an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or up to 100% sequence identity with SEQ ID NO: 100, and the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

226. The prime editor system of claim 225, wherein the dominant negative variant of MLH1 is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO:

213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.The prime editor system of claim 145, wherein the prime editor is PE2 of SEQ ID NO: 107 or an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or up to 100% sequence identity with SEQ ID NO: 107 and the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

227. The prime editor system of claim 202, wherein the prime editor is PE2 of SEQ ID NO: 107 or an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or up to 100% sequence identity with SEQ ID NO: 107, and the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

228. The prime editor system of claim 227, wherein the dominant negative variant of MLH1 is (a) MLH1 E34A (SEQ ID NO: 222), (b) MLH1 Δ756 (SEQ ID NO: 208), (c) MLH1 Δ754-756 (SEQ ID NO: 209), (d) MLH1 E34A Δ754-756 (SEQ ID NO: 210), (e) MLH1 1-335 (SEQ ID NO: 211), (f) MLH1 1-335 E34A (SEQ ID NO: 212), (g) MLH1 1-335 NLS^{SV40} (SEQ ID NO: 213), (h) MLH1 501-756 (SEQ ID NO: 215), (i) MLH1 501-753 (SEQ ID NO: 216), (j) MLH1 461-753 (SEQ ID NO: 218), or (k) NLS^{SV40} MLH1 501-753 (SEQ ID NO: 223), or a polypeptide comprising an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with any of SEQ ID NOs: 208-213, 215, 216, 218, 222, or 223.

229. The prime editor system of claim 202, wherein the prime editor is PE2 of SEQ ID NO: 107 or an amino acid sequence having at least 80%, at least 85%, at least 90%, at least 95%, at least 99%, or up to 100% sequence identity with SEQ ID NO: 107 and the inhibitor is a dominant negative variant of MLH1 that inhibits MLH1.

230. The prime editor system of claim 202, wherein the DNA polymerase is a reverse transcriptase.

231. The prime editor system of claim 230, wherein the reverse transcriptase is a retrovirus reverse transcriptase.

232. The prime editor system of claim 230, wherein the reverse transcriptase lacks RNase activity.

233. The prime editor system of claim 230, wherein the reverse transcriptase is a Moloney-Murine Leukemia Virus reverse transcriptase (MMLV-RT).

234. The prime editor system of claim 233, wherein the MMLV-RT comprises an amino acid sequence having at least 85% identity with a sequence selected from the group consisting of: SEQ ID NOs: 89 and 701-716.

235. The prime editor system of claim 202, wherein the napDNAbp is a CRISPR associated (Cas) nuclease.

236. The prime editor system of claim 235, wherein the napDNAbp comprises a Cas9 nuclease domain.

237. The prime editor system of claim 236, wherein the Cas9 nuclease domain is a nickase comprising a H840X substitution or a corresponding substitution as compared to a wild type *Streptococcus pyogenes* Cas9 as set forth in SEQ ID NO: 18, wherein X is any amino acid other than histidine.

238. The prime editor system of any one of claims 202-237, further comprising a pegRNA that is capable of complexing with the napDNAbp of the prime editor and programming the napDNAbp to bind a target DNA sequence.

239. A nucleic acid molecule encoding the prime editor system of any one of claims 202-238, or a component thereof.

240. A method for precisely installing a nucleotide edit of at least 15 bp in length in a double stranded target DNA sequence under conditions sufficient to evade the DNA mismatch repair pathway, the method comprising: contacting the double stranded target DNA sequence with a prime editor comprising a nucleic acid programmable DNA binding protein (napDNAbp), a DNA polymerase, and a prime editing guide RNA (PEgRNA), wherein the PEgRNA comprises a spacer that hybridizes to a first strand of the double stranded target DNA sequence, an extension arm that hybridizes to a second strand of the double stranded target DNA sequence, a DNA synthesis template comprising the nucleotide edit, and a gRNA core that interacts with the napDNAbp, and wherein the PEgRNA directs the prime editor to install the nucleotide edit in the double stranded target DNA sequence.

241. The method of claim 240, wherein the nucleotide edit is a deletion of at least 15 bp in length.

242. The method of claim 240, wherein the nucleotide edit is an insertion of at least 15 bp in length.

243. The method of claim 240, wherein the nucleotide edit is a least 16 bp, 17 bp, 18 bp, 19 bp, 20 bp, 21 bp, 22 bp, 23 bp, 24 bp, or 25 bp in length.

244. A method for editing a nucleic acid molecule by prime editing comprising: contacting a nucleic acid molecule with a prime editor and a pegRNA, thereby installing one or more modifications to the nucleic acid molecule at a target site, wherein the nucleic acid molecule is in a cell comprising a knockout of one or more genes involved in the DNA mismatch repair (MMR) pathway.

245. The method of claim 244, wherein the method further comprises contacting the nucleic acid molecule with a second strand nicking gRNA.

246. The method of claim 244, wherein the prime editing efficiency is increased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold,

at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold relative to a method performed in a cell that does not comprise a knockout of one or more genes involved in MMR.

247. The method of claim 244, wherein the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0 fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-

fold relative to a method performed in a cell that does not comprise a knockout of one or more genes involved in MMR.

248. The method of claim 244, wherein the one or more genes involved in MMR is selected from the group consisting of genes encoding the proteins MLH1, PMS2 (or MutL alpha), PMS1 (or MutL beta), MLH3 (or MutL gamma), MutS alpha (MSH2-MSH6), MutS beta (MSH2-MSH3), MSH2, MSH6, PCNA, RFC, EXO1, POLδ, and PCNA.

249. The method of claim 248, wherein the one or more genes is the gene encoding MLH1.

250. The method of claim 249, wherein MLH1 comprises an amino acid sequence of SEQ ID NO: 204, or an amino acid sequence having at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or up to and including 100% sequence identity with SEQ ID NO: 204.

251. The method of claim 244, wherein the prime editor comprises a napDNAbp and a polymerase.

252. The method of claim 251, wherein the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof.

253. The method of claim 251, wherein the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, Cas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute and optionally has a nickase activity.

254. The method of claim 251, wherein the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, or 99 (PEmax) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2, 4-67, or 99 (PEmax).

255. The method of claim 251, wherein the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 (i.e., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2.

256. The method of claim 251, wherein the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase.

257. The method of claim 251, wherein the polymerase is a reverse transcriptase.

258. The method of claim 257, wherein the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98.

259. The method of claim 251, wherein the napDNAbp and the polymerase of the prime editor are joined by a linker to form a fusion protein.

260. The method of claim 259, wherein the linker comprises an amino acid sequence of any one of SEQ ID NOs: 102 or 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102 or 118-131.

261. The method of claim 259, wherein the linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

262. The method of claim 244, wherein the one or more modifications to the nucleic acid molecule installed at the target site comprises one or more transitions, one or more transversions, one or more insertions, one or more deletions, or one more inversions, and optionally are less than 15 bp.

263. The method of claim 262, wherein the one or more transitions are selected from the group consisting of: (a) T to C; (b) A to G; (c) C to T; and (d) G to A.

264. The method of claim 262, wherein the one or more transversions are selected from the group consisting of: (a) T to A; (b) T to G; (c) C to G; (d) C to A; (e) A to T; (f) A to C; (g) G to C; and (h) G to T.

265. The method of claim 244, wherein the one or more modifications comprises changing (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair

to a C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair.

266. The method of claim 244, wherein the one or more modifications comprises an insertion or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides.

267. The method claim 244, wherein the one or more modifications comprises a correction to a disease-associated gene.

268. The method of claim 267, wherein the disease-associated gene is associated with a polygenic disorder selected from the group consisting of: heart disease; high blood pressure; Alzheimer's disease; arthritis; diabetes; cancer; and obesity.

269. A method for editing a nucleic acid molecule by prime editing comprising: contacting a nucleic acid molecule with a prime editor, a pegRNA, and an inhibitor of p53, thereby installing one or more modifications to the nucleic acid molecule at a target site.

270. The method of claim 269, wherein the method further comprises contacting the nucleic acid molecule with a second strand nicking gRNA.

271. The method of claim 269, wherein the prime editing efficiency is increased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least

37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of p53.

272. The method of claim 269, wherein the frequency of indel formation is decreased by at least 1.5-fold, at least 2.0-fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold, at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, at least 10.0-fold, at least 11-fold, at least 12-fold, at least 13-fold, at least 14-fold, at least 15-fold, at least 16-fold, at least 17-fold, at least 18-fold, at least 19-fold, at least 20-fold, at least 21-fold, at least 22-fold, at least 23-fold, at least 24-fold, at least 25-fold, at least 26-fold, at least 27-fold, at least 28-fold, at least 29-fold, at least 30-fold, at least 31-fold, at least 32-fold, at least 33-fold, at least 34-fold, at least 35-fold, at least 36-fold, at least 37-fold, at least 38-fold, at least 39-fold, at least 40-fold, at least 41-fold, at least 42-fold, at least 43-fold, at least 44-fold, at least 45-fold, at least 46-fold, at least 47-fold, at least 48-fold, at least 49-fold, at least 50-fold, at least 51-fold, at least 52-fold, at least 53-fold, at least 54-fold, at least 55-fold, at least 56-fold, at least 57-fold, at least 58-fold, at least 59-fold, at least 60-fold, at least 61-fold, at least 62-fold, at least 63-fold, at least 64-fold, at least 65-fold, at least 66-fold, at least 67-fold, at least 68-fold, at least 69-fold, at least 70-fold, at least 71-fold, at least 72-fold, at least 73-fold, at least 74-fold, or at least 75-fold, in the presence of the inhibitor of p53.

273. The method of claim 269, wherein the inhibitor of p53 is a protein.

274. The method of claim 273, wherein the protein is i53.

275. The method of claim 269, wherein the inhibitor of p53 is an antibody that inhibits the activity of p53.

276. The method of claim 269, wherein the inhibitor of p53 is a small molecule that inhibits the activity of p53.

277. The method of claim 269, wherein the inhibitor of p53 is a small interfering RNA (siRNA) or a small non-coding microRNA that inhibits the activity of p53.

278. The method of claim 269, wherein the prime editor comprises a napDNAbp and a polymerase.

279. The method of claim 278, wherein the napDNAbp is a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase domain or variant thereof.

280. The method of claim 278, wherein the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, Cas12b2, Cas13a, Cas12c, Cas12d, Cas12e, Cas12h, Cas12i, Cas12g, Cas12f (Cas14), Cas12f1, Cas12j (CasΦ), and Argonaute and optionally has a nickase activity.

281. The method of claim 278, wherein the napDNAbp comprises an amino acid sequence of any one of SEQ ID NOs: 2, 4-67, or 104, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 2, 4-67, or 104.

282. The method of claim 278, wherein the napDNAbp comprises an amino acid sequence of SEQ ID NO: 2 or SEQ ID NO: 37 (i.e., the napDNAbp of PE1 and PE2) or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with SEQ ID NO: 2 or SEQ ID NO: 37.

283. The method of claim 278, wherein the polymerase is a DNA-dependent DNA polymerase or an RNA-dependent DNA polymerase.

284. The method of claim 278, wherein the polymerase is a reverse transcriptase.

285. The method of claim 284, wherein the reverse transcriptase comprises an amino acid sequence of any one of SEQ ID NOs: 69-98 or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 69-98.

286. The method of claim 278, wherein the napDNAbp and the polymerase of the prime editor are joined by a linker to form a fusion protein.

287. The method of claim 286, wherein the linker comprises an amino acid sequence of any one of SEQ ID NOs: 102 or 118-131, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102 or 118-131.

288. The method of claim 286, wherein the linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 38, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

289. The method of claim 269, wherein the prime editor, the pegRNA, and the inhibitor of p53 are encoded on one or more DNA vectors.

290. The method of claim 289, wherein the one or more DNA vectors comprise AAV or lentivirus DNA vectors.

291. The method of claim 290, wherein the AAV vector is serotype 1, 2, 3, 4, 5, 6, 7, 8, 9, or 10.

292. The method of claim 286, wherein the prime editor as a fusion protein is further joined by a second linker to the inhibitor of p53.

293. The method of claim 292, wherein the second linker comprises a self-hydrolyzing linker.

294. The method of claim 292, wherein the second linker comprises an amino acid sequence of any one of SEQ ID NOs: 102, 118-131, or 233-236, or an amino acid sequence having at least an 80%, 85%, 90%, 95%, or 99% sequence identity with any one of SEQ ID NOs: 102, 118-131, or 233-236.

295. The method of claim 292, wherein the second linker is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, or 50 amino acids in length.

296. The method of claim 269, wherein the one or more modifications to the nucleic acid molecule installed at the target site comprises one or more transitions, one or more transversions, one or more insertions, one or more deletions, or one more inversions, and optionally wherein the one or more modifications are less than 15 bp.

297. The method of claim 296, wherein the one or more transitions are selected from the group consisting of: (a) T to C; (b) A to G; (c) C to T; and (d) G to A.

298. The method of claim 296, wherein the one or more transversions are selected from the group consisting of: (a) T to A; (b) T to G; (c) C to G; (d) C to A; (e) A to T; (f) A to C; (g) G to C; and (h) G to T.

299. The method of claim 269, wherein the one or more modifications comprises changing (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair to a C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair.

300. The method of claim 269, wherein the one or more modifications comprises an insertion or deletion of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 nucleotides.

301. The method claim 269, wherein the one or more modifications comprises a correction to a disease-associated gene.

302. The method of claim 301, wherein the disease-associated gene is associated with a polygenic disorder selected from the group consisting of: heart disease, high blood pressure, Alzheimer's disease, arthritis, diabetes, cancer, and obesity.

303. The method of claim 301, wherein the disease-associated gene is associated with a monogenic disorder selected from the group consisting of: Adenosine Deaminase (ADA) Deficiency, Alpha-1 Antitrypsin Deficiency, Cystic Fibrosis, Duchenne Muscular Dystrophy, Galactosemia, Hemochromatosis, Huntington's Disease, Maple Syrup Urine Disease, Marfan Syndrome, Neurofibromatosis Type 1, Pachyonychia Congenita, Phenylketonuria, Severe Combined Immunodeficiency, Sickle Cell Disease, Smith-Lemli-Opitz Syndrome, a trinucleotide repeat disorder, a prion disease, and Tay-Sachs Disease.

304. The method of any of the preceding claims, wherein the nucleic acid molecule is in a cell.

305. The method of claim 304, wherein the cell is a mammalian cell, a non-human primate cell, or a human cell.

306. The method of claim 304, wherein the cell is ex vivo.

307. The method of claim 304, wherein the cell is in a subject, optionally wherein the subject is human.

308. A method for treating a disease in a subject in need thereof, the method comprising administering to the subject: (i) a prime editor, (ii) a pegRNA, and (iii) an inhibitor of a DNA mismatch repair pathway, wherein the prime editor comprises a nucleic acid programmable DNA binding protein (napDNAbp) and a DNA polymerase,
wherein the pegRNA comprises a spacer sequence, a gRNA core, and an extension arm comprising a DNA synthesis template and a primer binding site (PBS),
wherein the spacer sequence comprises a region of complementarity to a target strand of a double stranded target DNA sequence in the subject,
wherein the gRNA core associates with the napDNAbp,
wherein the DNA synthesis template comprises a region of complementarity to the non-target strand of the double-stranded target DNA sequence and one or more nucleotide edits compared to the target strand double-stranded target DNA sequence;
wherein the primer binding site comprises a region of complementarity to a non-target strand of the double-stranded target DNA sequence,

wherein the prime editor and the pegRNA install the one or more nucleotide edits in the double stranded target DNA, wherein installation of the one or more nucleotide edits corrects one or more mutations in the double stranded target DNA associated with the disease, thereby treating the disease in the subject.

309. A method for treat a disease in a subject in need thereof, the method comprising administering to the subject: (i) a prime editor and (ii) a pegRNA wherein the prime editor comprises a nucleic acid programmable DNA binding protein (napDNAbp) and a DNA polymerase,

wherein the pegRNA comprises a spacer sequence, a gRNA core, and an extension arm comprising a DNA synthesis template and a primer binding site (PBS),
wherein the spacer sequence comprises a region of complementarity to a target strand of a double stranded target DNA sequence in the subject,
wherein the gRNA core associates with the napDNAbp,
wherein the DNA synthesis template comprises a region of complementarity to the non-target strand of the double-stranded target DNA sequence and comprises three or more consecutive nucleotide mismatches relative to the endogenous sequence of the double stranded target DNA sequence, wherein the three or more consecutive nucleotide mismatches comprise (i) an insertion, deletion, or substitution of x nucleotides that corrects a mutation (e.g. a disease associated mutation) and (ii) an insertion, deletion, or substitution of y nucleotides directly adjacent to the x nucleotides, wherein the insertion, deletion, or substitution of y nucleotides is a silent mutation, wherein (x+y) is an integer no less than 3, wherein y is an integer no less than 1, and wherein inclusion of the silent mutation(s) increases the efficiency, reduces unintended indel frequency, and/or improves editing outcome purity by prime editing,
wherein the primer binding site comprises a region of complementarity to a non-target strand of the double-stranded target DNA sequence,
wherein the prime editor and the pegRNA install the insertion, deletion, or substitution of x nucleotides in the double stranded target DNA, wherein installation insertion, deletion, or substitution of x nucleotides corrects one or more mutations in the double stranded target DNA associated with the disease, thereby treating the disease in the subject.

310. The method of claim 308 or 309, wherein the subject is a human.

311. A fusion protein comprising a nucleic acid programmable DNA binding protein (napDNAbp) domain and a domain comprising an RNA-dependent DNA polymerase activity, wherein the fusion protein comprises an amino acid sequence of SEQ ID NO: 99, or an amino acid sequence have at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% sequence identity with SEQ ID NO: 99

312. A fusion protein comprising a nucleic acid programmable DNA binding protein (napDNAbp) domain and a domain comprising an RNA-dependent DNA polymerase activity, wherein the napDNAbp comprises an amino acid sequence of SEQ ID NO: 104, or an amino acid sequence have at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% sequence identity with SEQ ID NO: 104.

313. A fusion protein comprising a nucleic acid programmable DNA binding protein (napDNAbp) domain and a domain comprising an RNA-dependent DNA polymerase activity, wherein the domain comprising an RNA-dependent DNA polymerase activity comprises an amino acid sequence of SEQ ID NO: 98, or an amino acid sequence have at least 80%, at least 85%, at least 90%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 100% sequence identity with SEQ ID NO: 98.

314. The fusion protein of any of claims 311-313, further comprising a linker that joins the napDNAbp and the domain comprising the RNA-dependent DNA polymerase activity.

315. The fusion protein of claim 314, wherein the linker comprises SEQ ID NO: 105.

316. The fusion protein of claim 311, wherein the napDNAbp is a Cas9 nickase.

317. The fusion protein of claim 316, wherein the Cas9 nickase comprises an H840A substitution and at least one substitution at R221 or N394 relative to SEQ ID NO: 37.

318. A complex comprising a fusion protein of any one of claims 311-317 and a PEgRNA, wherein the PEgRNA directs the fusion protein to a target DNA sequence for prime editing.

319. The complex of claim 318, wherein the PEgRNA comprises a guide RNA and a nucleic acid extension arm at the 3′ or 5′ end of the guide RNA.

320. The complex of claim 319, wherein the PEgRNA is capable of binding to a napDNAbp and directing the napDNAbp to the target DNA sequence.

321. A polynucleotide encoding the fusion protein of any of claims 311-317.

322. A vector comprising the polynucleotide of claim 321, wherein expression of the fusion protein is under the control of a promoter.

323. The vector of claim 322, wherein the promoter is a U6 promoter.

324. A cell comprising the fusion protein of any of claims 311-317 and a PEgRNA bound to the napDNAbp of the fusion protein.

325. A cell comprising a complex of any one of claims 318-320.

326. A pharmaceutical composition comprising: (i) a fusion protein of any of claims 311-317, the complex of claims 318-320, the polynucleotide of claim 321, or the vector of claims 322-323; and (ii) a pharmaceutically acceptable excipient.

327. A method for editing a nucleic acid molecule by prime editing comprising: contacting a nucleic acid molecule with a modified prime editor of any one of claims 311-317 and a pegRNA, thereby installing one or more modifications to the nucleic acid molecule at a target site.

328. The method of claim 327, wherein the method further comprises contacting the nucleic acid molecule with a second strand nicking gRNA.

329. The method of claim 327, wherein the prime editing efficiency is increased by at least 1.5-fold, at least 2.0 fold, at least 2.5-fold, at least 3.0-fold, at least 3.5-fold, at least 4.0-fold,

at least 4.5-fold, at least 5.0-fold, at least 5.5-fold, at least 6.0-fold, at least 6.5-fold, at least 7.0-fold, at least 7.5-fold, at least 8.0-fold, at least 8.5-fold, at least 9.0-fold, at least 9.5-fold, or at least 10.0 fold relative to prime editing with PE2.