ADENOSINE NUCLEOBASE EDITORS AND USES THEREOF

BACKGROUND OF THE INVENTION

[0001] Targeted editing of nucleic acid sequences, for example, the targeted cleavage or the targeted introduction of a specific modification into genomic DNA, is a highly promising approach for the study of gene function and also has the potential to provide new therapies for human genetic diseases. Since many genetic diseases in principle can be treated by effecting a specific nucleotide change at a specific location in the genome (for example, an A to G or a T to C change in a specific codon of a gene associated with a disease), the development of a programmable way to achieve such precise gene editing represents both a powerful new research tool, as well as a potential new approach to gene editing-based therapeutics.

SUMMARY OF THE INVENTION

[0002] Provided herein are compositions, kits, and methods of modifying a polynucleotide (*e.g.*, DNA) using an adenosine deaminase and a nucleic acid programmable DNA binding protein (*e.g.*, Cas9) Some aspects of the disclosure provide nucleobase editing proteins which catalyze hydrolytic deamination of adenosine (forming inosine, which base pairs like guanine (G)) in the context of DNA. There are no known naturally occurring adenosine deaminases that act on DNA. Instead, known adenosine deaminases act on RNA (*e.g.*, tRNA or mRNA). To overcome this drawback, the first deoxyadenosine deaminases were evolved to accept DNA substrates and deaminate deoxyadenosine (dA) to deoxyinosine. The adenosine deaminase acting on tRNA (ADAT) from *Escherichia coli* (TadA, for tRNA adenosine deaminase A), was covalently fused to a dCas9 domain, and libraries of this fusion were assembled containing mutations in the deaminase portion of the construct. It should be appreciated that *E. coli* TadA (ecTadA) deaminases also include truncations of ecTadA. For example, truncations (*e.g.*, N-terminal truncations) of a full length ecTadA (SEQ ID NO: 84), such as the N-terminally truncated ecTadA set forth in SEQ ID NO: 1 are provided herein for use in the present invention. Further, it was found that other adenosine deaminase mutants, such as *S. aureus* TadA mutants, were capable of deaminating adenosine. Without wishing to be bound by any particular theory, truncations of adenosine deaminases (*e.g.*, ecTadA) may have desired solubility and/or expression properties as compared to their full-length counterparts.

[0003] Mutations in the deaminase domain of nucleobase editing proteins were made by evolving adenosine deaminases. Productive variants were identified via selection for A to G reversion at the codon of an active-site His in the acetyl-transferase gene of chloramphenicol (encoded on a co-transformed selection plasmid). A first round of evolution yielded an ecTadA variant, ecTadA D108X (X=G, V, or N), capable of converting A to G in DNA. In some embodiments, the ecTadA variant comprises a D108A mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase. The first round of evolution also yielded an ecTadA variant, ecTadA A106V. A subsequent round of evolution resulted in another variant, ecTadA D108N_E155X (X=G, V, or D), which *E. coli* survive in the presence of high concentrations of chloramphenicol. Additional variants were identified by evolving ecTadA. For example, ecTadA variants that are capable of deaminating adenosine in DNA include one or more of the following mutations D108N, A106V, D147, E155V, L84F, H123Y, and I157F of SEQ ID NO: 1. It should be appreciated however, that homologous mutations may be made in other adenosine deaminases to generate variants that are capable of deaminating adenosine in DNA. Additional rounds of evolution provided further ecTadA variants. For example, additional ecTadA variants are shown in Figures 11, 16, 97, 104-106, 125-128, 115 and Table 4.

[0004] In the examples provided herein, exemplary nucleobase editors having the general structure evolved ecTadA(D108X; X=G, V, or N)-XTEN-nCas9, catalyzed A to G transition mutations in cells such as eukaryotic cells (*e.g.*, Hek293T mammalian cells). In other examples exemplary nucleobase editors contain two ecTadA domains and a nucleic acid programmable DNA binding protein (napDNAbp). For example nucleobase editors may have the general structure ecTadA(D108N)-ecTadA(D108N)-nCas9. Additional examples of nucleobase editors containing ecTadA variants provided herein demonstrate an improvement in performance of the nucleobase editors in mammalian cells. For example, certain adenosine base editors include ecTadA having D108X, where X=G, V, or N, and/or E155X, where X=B, V, or D mutations in ecTadA as set forth in SEQ ID NO: 1 or another adenine deaminase. In certain embodiments mutants, nucleobase editors are covalently fused to catalytically dead alkyl adenosine gylcosylase (AAG), which may protect the edited inosine from base excision repair (or other DNA repair systems) until the T on the opposite strand is changed to a C, for example, through mismatch repair (or other DNA repair systems). Once the base opposite the inosine is changed to a C, then the inosine may be changed to a G irreversibly and permanently through cellular DNA repair processes, resulting in a permanent change from an A:T base pair to a G:C base pair.

[0005] Without wishing to be bound by any particular theory, the adenosine nucleobase editors described herein work by using ecTadA variants to deaminate A bases in DNA, causing A to G mutations via inosine formation. Inosine preferentially hydrogen bonds with C, resulting in A to G mutation during DNA replication. When covalently tethered to Cas9 (or another nucleic acid programmable DNA binding protein), the adenosine deaminase (*e.g.,* ecTadA) is localized to a gene of interest and catalyzes A to G mutations in the ssDNA substrate. This editor can be used to target and revert single nucleotide polymorphisms (SNPs) in disease-relevant genes, which require A to G reversion. This editor can also be used to target and revert single nucleotide polymorphisms (SNPs) in disease-relevant genes, which require T to C reversion by mutating the A, opposite of the T, to a G. The T may then be replaced with a C, for example by base excision repair mechanisms, or may be changed in subsequent rounds of DNA replication.

[0006] Some aspects of the disclosure relate to the discovery that engineered (*e.g.,* evolved) adenosine deaminases are capable of deaminating adenosine in a deoxyribonucleic acid (DNA) substrate. In some embodiments, the disclosure provides such adenosine deaminases. In some embodiments, the adenosine deaminases provided herein are capable of deaminating an adenosine in a DNA molecule. Other aspects of the disclosure provide fusion proteins comprising a Cas9 domain and an adenosine deaminase domain, for example, an engineered deaminase domain capable of deaminating an adenosine in DNA. In some embodiments, the fusion protein comprises one or more of a nuclear localization sequence (NLS), an inhibitor of inosine base excision repair (*e.g.,* dISN), and/or a linker.

[0007] In some aspects, the disclosure provides an adenosine deaminase capable of deaminating an adenosine in a deoxyribonucleic acid (DNA) substrate. In some embodiments, the adenosine deaminase is from a bacterium, for example, *E. coli* or *S. aureus*. In some embodiments, the adenosine deaminase is a TadA deaminase. In some embodiments, the TadA deaminase is an *E. coli* TadA deaminase (ecTadA). In some embodiments, the adenosine deaminase comprises a D108X mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, wherein X is any amino acid other than the amino acid found in the wild-type protein. In some embodiments, X is G, N, V, A, or Y.

[0008] In some embodiments, the adenosine deaminase comprises a E155X mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, wherein X is any amino acid other than the amino acid found in the wild-type protein. In some embodiments,

X is D, G, or V. It should be appreciated that the adenosine deaminases provided herein may contain one or more of the mutations provided herein in any combination.

[0009] Some aspects of the disclosure provide a fusion protein comprising: (i) a Cas9 domain, and (ii) an adenosine deaminase, such as any of the adenosine deaminases provided herein. In some embodiments, the Cas9 domain of the fusion protein is a nuclease dead Cas9 (dCas9), a Cas9 nickase (nCas9), or a nuclease active Cas9. In some embodiments, the fusion protein further comprises an inhibitor of inosine base excision repair, for example a dISN or a single stranded DNA binding protein. In some embodiments, the fusion protein comprises one or more linkers used to attach an adenine deaminase (*e.g.,* ecTadA) to a nucleic acid programmable DNA binding protein (*e.g.,* Cas9). In some embodiments, the fusion protein comprises one or more nuclear localization sequences (NLS).

[0010] The summary above is meant to illustrate, in a non-limiting manner, some of the embodiments, advantages, features, and uses of the technology disclosed herein. Other embodiments, advantages, features, and uses of the technology disclosed herein will be apparent from the Detailed Description, the Drawings, the Examples, and the Claims.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] Figure 1 shows high throughput screen results with various deaminases. APOBEC (BE3) is the positive control; ADAR acts on mRNA, ADA acts on deoxyadenosine, and ADAT acts on tRNA. The untreated group is the negative control. The sequence corresponds to SEQ ID: 45.

[0012] Figure 2 is a schematic of a deamination selection plasmid.

[0013] Figure 3 shows a serial dilution of the selection plasmid in S1030 cells plated on increasing concentrations of chloramphenicol.

[0014] Figure 4 shows the validation of chloramphenicol selection with a rAPOBEC1-XTEN-dCas9 construct as a positive control. The sequences from top to bottom correspond to SEQ ID NOs: 95 (the nucleotide sequence), 96 (the amino acid sequence), 97 (the nucleotide sequence), 98 (the amino acid sequence), 95 (the nucleotide sequence) and 99 (the truncated nucleotide sequence).

[0015] Figure 5 is a schematic of a deaminase-XTEN-dCas9 construct.

[0016] Figure 6 shows the sequencing results from the first round of the TadA-XTEN-dCas9 library.

[0017] Figure 7 shows the sequence of a selection plasmid; an A to G reversion was observed. The sequences from top to bottom correspond to SEQ ID NOs: 100 (the nucleotide

sequence), 101 (the amino acid sequence), 102 (the nucleotide sequence), 103 (the amino acid sequence), 104 (the nucleotide sequence), and 100 (the nucleotide sequence).

[0018] Figure 8 shows the results of deaminase sequencing, illustrating the convergence at residue D108. The sequences correspond to SEQ ID NOs: 589-607 from top to bottom.

[0019] Figure 9 shows the *E. coli* TadA crystal structure. Note that D119 in the figure corresponds to D108, as the residue numbering isoffset in the figure.

[0020] Figure 10 shows the crystal structure of TadA (in *S. aureus*) tRNA and an alignment of with TadA from *E. coli.* The sequences from top to bottom correspond to SEQ ID NOs: 105-107.

[0021] Figure 11 shows results from the isolation and challenge of individual constructs from ecTadA evolution.

[0022] Figure 12 shows the colony forming units (C.F.U.) of various constructs challenged on increasing concentrations of chloramphenicol. The construct numbers correspond to those listed in Figure 11.

[0023] Figure 13 shows data from the second round of evolution from the constructs containing the D108N mutation. The sequences from top to bottom correspond to SEQ ID NOs: 608-623.

[0024] Figure 14 shows A to G editing in mammalian cells. The sequence corresponds to SEQ ID NO: 41.

[0025] Figure 15 is a schematic showing the development of ABE.

[0026] Figure 16 is a table showing the results of clones assayed after second round evolution. Columns 1, 8, and 10 represent mutations from the first round evolution. Columns 11 and 14 represent the consensus mutations from second round evolution.

[0027] Figure 17 shows the results of individual clone antibiotic challenge assays. The identity of the construct numbers correspond to the pNMG clone numbers from Figure 16.

[0028] Figure 18 show schematic representations of new constructs that were developed. New constructs include UGI, AAG*E125Q, and EndoV*D35A domains.

[0029] Figure 19 shows the transfection of constructs into mammalian cells containing single or double mutations in ecTadA. The sequence corresponds to SEQ ID NO: 41.

[0030] Figure 20 shows the transfection of constructs with the addition of UGI to adenosine nucleobase editor (ABE) (D108N). The sequence corresponds to SEQ ID NO: 41.

[0031] Figure 21 shows that ABE operates best on 1 of 6 genomic sites tested. The sequence corresponds to SEQ ID NO: 46.

[0032] Figure 22 shows that the Hek-3 site also has lower editing relative to the Hek-2 site editing at position 8 of the protospacer. The sequence corresponds to SEQ ID NO: 42.

[0033] Figure 23 shows inactive C-terminal Cas9 fusions of ecTadA for constructs pNMG-164 through pNMG-173. The sequence corresponds to SEQ ID NO: 41.

[0034] Figure 24 shows inactive C-terminal Cas9 fusions of ecTadA for pNMG-174 through pNMG-177. The sequence corresponds to SEQ ID NO: 41.

[0035] Figure 25 shows the editing results from ecTadA nucleobase editors (pNMG-143, pNMG-144, pNMG-164, and pNMG-177). The sequence corresponds to SEQ ID NO: 41.

[0036] Figure 26 shows the editing results from ecTadA nucleobase editors (pNMG-164, pNMG-177, pNMG-178, pNMG-179, and pNMG-180). The sequence corresponds to SEQ ID NO: 41.

[0037] Figure 27 shows the results of editing at the Hek-3 site. The sequence corresponds to SEQ ID NO: 42.

[0038] Figure 28 shows the results of editing at the Hek-2 site. The sequence corresponds to SEQ ID NO: 41.

[0039] Figure 29 shows the results of editing at the Hek-3 site. The sequence corresponds to SEQ ID NO: 42.

[0040] Figure 30 shows the results of editing at the Hek-4 site. The sequence corresponds to SEQ ID NO: 43.

[0041] Figure 31 shows the results of editing at the RNF-2 site. The sequence corresponds to SEQ ID NO: 44.

[0042] Figure 32 shows the results of editing at the FANCF site. The sequence corresponds to SEQ ID NO: 45.

[0043] Figure 33 shows the results of editing at the EMX-1 site. The sequence corresponds to SEQ ID NO: 46.

[0044] Figure 34 shows the results of C-terminal fusion at the Hek-2 site. The sequence corresponds to SEQ ID NO: 41.

[0045] Figure 35 shows the results of C-terminal fusion at the Hek-3 site. The sequence corresponds to SEQ ID NO: 42.

[0046] Figure 36 shows the results of C-terminal fusion at the Hek-4 site. The sequence corresponds to SEQ ID NO: 43.

[0047] Figure 37 shows the results of C-terminal fusion at the EMX-1 site. The sequence corresponds to SEQ ID NO: 46.

[0048] Figure 38 shows the results of C-terminal fusion at the RNF-2 site. The sequence corresponds to SEQ ID NO: 44.

[0049] Figure 39 shows the results of C-terminal fusion at the FANCF site. The sequence corresponds to SEQ ID NO: 45.

[0050] Figure 40 shows the results of transfection at the Hek-2 site. The sequence corresponds to SEQ ID NO: 41.

[0051] Figure 41 shows the results of transfection at the Hek-3 site. The sequence corresponds to SEQ ID NO: 42.

[0052] Figure 42 shows the results of transfection at the RNF-2 site. The sequence corresponds to SEQ ID NO: 44.

[0053] Figure 43 shows the results of transfection at the Hek-4 site. The sequence corresponds to SEQ ID NO: 43.

[0054] Figure 44 shows the results of transfection at the EMX-1 site. The sequence corresponds to SEQ ID NO: 46.

[0055] Figure 45 shows the results of transfection at the FANCF site. The sequence corresponds to SEQ ID NO: 45.

[0056] Figure 46 shows deaminase editing of sgRNA.

[0057] Figure 47 shows constructs developed for fusions at various sites.

[0058] Figure 48 shows indel rates for different fusions at various sites.

[0059] Figure 49 shows the protospacer and PAM sequences of base editing sites set forth in SEQ ID NOs: 46, 45, 6, 42, 43, and 468from top to bottom, respectively.

[0060] Figure 50 shows constructs developed for fusions at various sites using further mutated D108 residue.

[0061] Figure 51 shows the protospacer and PAM sequences of base editing sites set forth in SEQ ID NOs: 6, 46, and 42 from top to bottom, respectively.

[0062] Figure 52 shows the results of using mutated D108 residues to cause deaminase to reject RNA as a substrate and change the editing outcome.

[0063] Figure 53 shows the results of using mutated D108 residues to cause deaminase to reject RNA as a substrate and change the editing outcome.

[0064] Figure 54 shows constructs developed for fusions at various sites.

[0065] Figure 55 shows the protospacer and PAM sequences of base editing sites set forth in SEQ ID NOs: 6, 358, 359from top to bottom, respectively.

[0066] Figure 56 shows the results of ABE on HEK site 2.

[0067] Figure 57 shows the results of ABE on HEK site 2.

[0068] Figure 58 shows constructs developed for fusions at various sites using various linker lengths.

[0069] Figure 59 shows the importance of linker length on base editing function.

[0070] Figure 60 shows the importance of linker length on base editing function.

[0071] Figure 61 is a schematic showing the dimerization of deaminase.

[0072] Figure 62 shows constructs developed for fusions at various sites using various linker lengths.

[0073] Figure 63 shows the current editor architecture (top panel), the *in trans* dimerization (bottom panel, left), and the *in cis* dimerization (bottom panel, right).

[0074] Figure 64 shows dimerization results from base editing.

[0075] Figure 65 shows dimerization results from base editing.

[0076] Figure 66 shows dimerization results from base editing.

[0077] Figure 67 shows constructs developed for fusions at various sgRNA sites.

[0078] Figure 68 shows the evolution of ABE editor against new selection sequences. The sequences from top to bottom and left to right correspond to SEQ ID NOs: 707-719, respectively.

[0079] Figure 69 shows the current editor targeting Q4 stop site. The sequences from top to bottom correspond to SEQ ID NOs: 624-628.

[0080] Figure 70 shows the current editor targeting W15 stop site. The sequences correspond to SEQ ID NOs: 629-633 from top to bottom respectively.

[0081] Figure 71 shows a HEK293 site 2 sequence. The sequence corresponds to SEQ ID NO: 360.

[0082] Figure 72 shows the results of the first run with various edTadA mutations using the sequence of Figure 71.

[0083] Figure 73 shows the results of the second run with various edTadA mutations using the sequence of Figure 71.

[0084] Figure 74 shows a FANCF sequence. The sequence corresponds to SEQ ID NO: 45.

[0085] Figure 75 shows the results of the second run using various edTadA mutations and the sequence of Figure 74.

[0086] Figure 76 shows the results of mutated D108 on all sites.

[0087] Figure 77 shows in trans data from previous run (left panel) and the mut-mut fusions hindered by super long linkers.

[0088] Figure 78 shows the results of tethering mutTadA to ABE.

[0089] Figure 79 shows the constructs of all inhibitors tested.

[0090] Figure 80 shows the constructs used when tethering AAG to ABE.

[0091] Figure 81 is a schematic showing the tethering of AAG to ABE.

[0092] Figure 82 shows the results of tethering AAG to ABE.

[0093] Figure 83 shows the constructs used when tethering AAG to ABE with an N-terminus of TadA.

[0094] Figure 84 is a schematic showing the tethering of AAG to ABE with an N-terminus of TadA.

[0095] Figure 85 shows the results of tethering AAG to ABE.

[0096] Figure 86 shows the constructs used when tethering EndoV to ABE.

[0097] Figure 87 is a schematic showing the tethering EndoV to ABE.

[0098] Figure 88 shows the results of tethering EndoV to ABE.

[0099] Figure 89 shows the constructs used when tethering UGI to ABE.

[00100] Figure 90 shows the results of tethering UGI to the end of ABE.

[00101] Figure 91 shows the results of various inhibitors increasing A to G editing.

[00102] Figure 92 shows a sequence alignment of prokaryotic TadA amino acid sequences. The sequences correspond to SEQ ID NOs: 634-657 from top to bottom respectively.

[00103] Figure 93 shows a schematic of the relative sequence identity analysis of TadA amino acid sequences.

[00104] Figure 94 shows a schematic representation of an exemplary adenosine base editing process.

[00105] Figure 95 shows a schematic representation of an exemplary adenosine base editor, which deaminates adenosine to inosine.

[00106] Figure 96 shows a schematic of an exemplary base-editing selection plasmid.

[00107] Figure 97 shows a list of clones including identified mutations in ecTadA.

[00108] Figure 98 shows an exemplary sequencing analysis of a selection plasmid from surviving colonies. The sequences correspond to SEQ ID NOs: 658-661 from top to bottom respectively.

[00109] Figure 99 shows a schematic of exemplary adenosine base editors from a third round of evolution.

[00110] Figure 100 shows the percentage of A to G conversions in Hek293T cells.

[00111] Figure 101 shows a schematic of an exemplary base-editing selection plasmid.

[00112] **Figure 102** shows a schematic representation of the verdine crystal structure of *S. aureus* TadA. The *S. aureus* TadA, a homolog of ecTadA, is shown with its tRNA substrate co-crystalized. Red arrows are the H-bond contacts with the various nucleic acids in the tRNA substrate. SeeLosey, H.C., et al., "Crystal structure of Staphylococcus sureus tRNA adenosine deaminase tadA in complex with RNA", *Nature Struct*. Mol. Biol. 2, 153-159 (2006).

[00113] **Figure 103** shows a schematic of a construct containing ecTadA_2.2 and dCas9, identifying mutated ecTadA residues.

[00114] **Figure 104** shows results of ecTadA evolution (evolution #4) at sites E25 and R26.

[00115] **Figure 105** shows results of ecTadA evolution (evolution #4) at site R107.

[00116] **Figure 106** shows results of ecTadA evolution (evolution #4) at sites A142 and A143.

[00117] **Figure 107** shows an exemplary sequencing analysis of a selection plasmid from surviving colonies. The sequences correspond to SEQ ID NO: 662-671 from top to bottom respectively.

[00118] **Figure 108** shows a summary of results of editing at the Hek-2 site. The Hek-2 sequence provided in the figure represents the reverse complement of SEQ ID NO: 41, which is the DNA strand where A to G editing takes place. The sequence corresponds to SEQ ID ID: 6.

[00119] **Figure 109** shows a summary of results of editing at the Hek2-3 site. The sequence corresponds to SEQ ID NO: 363.

[00120] **Figure 110** shows a summary of results of editing at the Hek2-6 site. The sequence corresponds to SEQ ID NO: 364.

[00121] **Figure 111** shows a summary of results of editing at the Hek2-7 site. The Hek2-7 sequence provided in the figure represents the reverse complement of the DNA strand where A to G editing takes place. The sequence corresponds to SEQ ID NO: 365.

[00122] **Figure 112** shows a summary of results of editing at the Hek2-10 site. The sequence corresponds to SEQ ID NO: 366.

[00123] **Figure 113** shows a summary of results of editing at the Hek-3 site. The sequence corresponds to SEQ ID NO: 42.

[00124] **Figure 114** shows a summary of results of editing at the FANCF site. The sequence corresponds to SEQ ID NO: 45.

[00125] **Figure 115** shows a summary of results of editing at the Hek-2 site. The sequence corresponds to SEQ ID NO: 367.

[00126] **Figure 116** shows a summary of results of editing at the Hek2-2 site. The sequence corresponds to SEQ ID NO: 368.

[00127] **Figure 117** shows a summary of results of editing at the Hek2-3 site. The sequence corresponds to SEQ ID NO: 363.

[00128] **Figure 118** shows a summary of results of editing at the Hek2-6 site. The sequence corresponds to SEQ ID NO: 364.

[00129] **Figure 119** shows a summary of results of editing at the Hek2-7 site. The sequence corresponds to SEQ ID NO: 365.

[00130] **Figure 120** shows a summary of results of editing at the Hek2-10 site. The sequence corresponds to SEQ ID NO: 366.

[00131] **Figure 121** shows a summary of results of editing at the Hek-3 site. The sequence corresponds to SEQ ID NO: 42.

[00132] **Figure 122** shows a summary of results of editing at the FANCF site. The sequence corresponds to SEQ ID NO: 45.

[00133] **Figure 123** shows the results of ecTadA evolution (evolution #4) at HEK2, HEK2-2, HEK2-3, HEK2-6, HEK2-7, and HEK2-10 sites. The constructs used were pNMG-370 (evolution #2), pNMG-371 (evolution #3), and pNMG 382-389 (evolution #4). The sequences correspond to SEQ ID NOs: 7, 368, 363, 364, 369, and 370 from top to bottom., respectively.

[00134] **Figure 124** shows a schematic of a construct containing ecTadA and dCas9 used for ecTadA evolution (evolution #5).

[00135] **Figure 125** is a table showing the results of clones assayed after fifth round evolution (128 ug/mL chlor, 7h).

[00136] **Figures 126A to 136B** are tables showing the results of sub-cloned and re-transformed clones assayed after fifth round under varying conditions.

[00137] **Figure 127** is a table showing the results of amplicons from spectinomycin selection clones assayed after fifth round evolution.

[00138] **Figure 128** is a table showing the results of clones assayed after fifth round evolution.

[00139] **Figure 129** shows a summary of results of editing at the Hek-2 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349.

As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The sequence corresponds to SEQ ID NO: 6.

[00140] **Figure 130** shows a summary of results of editing at the Hek2-1site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The Hek2-1 sequence provided in the figure represents the DNA strand where A to G editing takes place. The sequence corresponds to SEQ ID NO: 465.

[00141] **Figure 131** shows a summary of results of editing at the Hek2-2 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The sequence corresponds to SEQ ID NO: 368.

[00142] **Figure 132** shows a summary of results of editing at the Hek2-3 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The sequence corresponds to SEQ ID NO: 363.

[00143] **Figure 133** shows a summary of results of editing at the Hek2-4 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The Hek2-4 sequence provided in the figure represents the DNA strand where A to G editing takes place. The sequence corresponds to SEQ ID NO: 466.

[00144] **Figure 134** shows a summary of results of editing at the Hek2-6 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The sequence corresponds to SEQ ID NO: 364.

[00145] **Figure 135** shows a summary of results of editing at the Hek2-9 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The Hek2-9 sequence provided in the figure represents the DNA strand where A to G editing takes place. The sequence corresponds to SEQ ID NO: 467.

[00146] **Figure 136** shows a summary of results of editing at the Hek2-10 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-

349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The Hek2-10 sequence provided in the figure represents the DNA strand where A to G editing takes place. The sequence corresponds to SEQ ID NO: 370.

[00147] **Figure 137** shows a summary of results of editing at the Hek3 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The sequence corresponds to SEQ ID NO: 42.

[00148] **Figure 138** shows a summary of results of editing at the RNF2 site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The sequence corresponds to SEQ ID NO: 468.

[00149] **Figure 139** shows a summary of results of editing at the FANCF site using base editors that contain an engineered *S. aureus* TadA (saTadA), which include pNMG-346-349. As a comparison, results of editors that contain an engineered *E. coli* TadA (ecTadA), which include pNMG-339-341, are shown. The sequence corresponds to SEQ ID NO: 45.

[00150] **Figure 140** shows various schematic representations of adenosine base editor (ABE) constructs. The identity of the editors e.g., "pNMG-367" is indicated in Table 4. The following mutations are abbreviated as follows: ecTadA1 (A106V D108N), ecTadA2 (A106V D108N D147Y E155V), ecTadA3 (ecTadA2 + L84F H123Y I156F), ecTadA3+ (ecTadA3 + A142N), ecTadA5a1 (ecTadA3 + H36L R51L S146C K157N), ecTadA5a3 (ecTadA3 + N37S K161T), ecTadA5a11 (ecTadA3 + R51L S146C K157N K161T), ecTadA5a12 (ecTadA3 + S146C K161T), ecTadA5a14 (ecTadA3 + RS146C K157N K160E), and ecTadA5a1+ (ecTadA5a1 + A142N), ecTadA5a9 (ecTadA3 + S146R K161T). Heterodimers of the top three ABE 5a constructs were made and then tested relative to homodimers. The heterodimer version of the ABE editor typically performs better than the corresponding homodimeric construct. Both homodimeric and heterodimeric constructs are shown in Figure 140.

[00151] **Figure 141** shows editing results for various ABE constructs. The ABE plasmid # refers to pNMG number as indicated in Table 4. For example 367 refers to construct pNMG-367 in Table 4. The sequences correspond to SEQ ID NOs: 469 (pNMG-466), 470 (pNMG-467), 471 (pNMG-469), 472 (pNMG-470), 473 (pNMG-501), 474 (pNMG-509), and 475 (pNMG-502) from top to bottom, respectively.

[00152] **Figure 142** shows editing results for various ABE constructs at specific sites. The numbers on the top row indicate the pNMG number as indicated in Table 4. For example 107 refers to construct pNMG-107 in Table 4. In certain contexts, homodimer constructs have been shown to work better than a hetero dimer construct and *vice versa* (see for example construct 371 which is a homodimer versus construct 476 which is a heterodimer). Schematics for these ABE constructs are shown in Figure 140, and the construct architecture is shown in Table 4. The sequences correspond to SEQ ID NOs: 478, 478, 514, 516, 516, 520, 520, 521, 521, and 509 from top to bottom, respectively.

[00153] **Figure 143** shows the percentage of indels formed for ABE constructs from Figure 142.

[00154] **Figure 144** shows editing results for various ABE constructs at specific sites. The identity of the constructs are shown in the top row and refer to the pNMG reference number of Table 4. The results in Figure 144 indicate that adding ecTadA monomer to ABE construct may not improve editing. However, adding a long linker between monomers may help editing at some sites (see, for example, the editing results for sgRNA constructs 285b versus 277 at sites 502, 505, 507). The identity of the sgRNA constructs is shown in Table 8 Schematics for these ABE constructs are shown in **Figure 140**. The sequences correspond to SEQ ID NOs: 478, 480, 480, 514, 517, 517, 517, 517, 519, and 521 from top to bottom, respectively.

[00155] **Figure 145** shows results for ABE constructs at all NAN sites, where the target A is at position 5 of the Protospacer and PAM sequencs. The identity of the ABE constructs, shown in the top row refers to the pNMG reference number in Table 4. The number values represent the % of target A residues that were edited (*e.g.,* % editing efficiency). The sequences correspond to SEQ ID NOs: 537-552 from top to bottom, respectively.

[00156] **Figure 146** shows A to G editing percent at the Hek2 site for various ABE constructs as referenced by their reference pNMG number in Table 4.

[00157] **Figure 147** shows evolution round #5b evolution results. The number values represent the % of A to G editing for the indicated sites. The sequences from top to bottom correspond to SEQ ID NOs: 7, 465, 368, 363, 364, and 370 from top to bottom, respectively.

[00158] **Figure 148** shows editing results for various ABE constructs which were obtained from different rounds of evolution (*e.g.,* evo3). The generic schematic for the ABE constructs is also shown. The identity of the sgRNA, as indicated in Table 8, and the identity of the base editors (pNMG reference), as indicated in Table 4, are shown. The number values

represent the % of A to G editing for the indicated sites. The sequences correspond to SEQ ID NOs: 478, 503, 506, 521, 513, 505, 507, and 509 from top to bottom, respectively.

[00159] **Figure 149** shows examination of the ABE constructs at genomic sites other than the Hek-2 sequence. The Hek-2 site (sgRNA 299) is represented by the asterisk. The identity of the sgRNA is indicated in Table 8. The sequences correspond to SEQ ID NOs: 478, 514, 516, 517, 517, 517, 517, 519, 520, 529, 521 from top to bottom, respectively.

[00160] **Figure 150** shows a schematic of the DNA shuffling experiment using nucleotide exchange and excision technology (NExT), which is referred to as ABE evolution #6. The goal of this approach was to assemble a more efficient editor and remove potential epistatic mutations. DNA shuffling of constructs from various evolutions were used to optimize for desired mutations and eliminate mutations that negatively affect editing efficiencies and/or protein stability.

[00161] **Figure 151** shows a schematic for DNA Shuffle (NeXT). The spect target sequence is 5'-CAATG**A**TGACTTCTACAGCG-3' (SEQ ID NO: 444) and the chlor target sequence is 5'-TACGGCGT**A**GTGCACCTGGA-3' (SEQ ID NO: 441).

[00162] **Figure 152** shows the sequence identity of clones from evolution #6 surviving on spect only (non-YAC target). The mutations indicated are relative to ecTadA (SEQ ID NO: 1).

[00163] **Figure 153** shows evolution # 6.2 which refers to the enrichment of clones from evolution #6. The mutations indicated are relative to ecTadA (SEQ ID NO: 1). A142N is present in almost all clones sequenced and the Pro48 mutation is also abundant. The clones were selected against "GAT" in the spectinomycin site. The selection target sequence was 5'-CAATG**A**TGACTTCTACAGCG-3' (SEQ ID NO: 444).

[00164] **Figure 154** shows schematic representations of ABE 6 constructs. 8 new constructs in total were developed. Mutations from the top 2 highest frequency amplicons in Evo #6 were used in each of the four architectures.

[00165] **Figure 155** shows data harvesting for ABE: step 1 – transfection + HTS of key intermediates at 6 genomic sites, n = 3. The transfection was performed with 750 ng ABE + 250 ng gRNA and incubated for 5 days before the genomic DNA was extracted to perform HTS. The identity of each of the ABE constructs is indicated by the pNMG reference number as shown in Table 4. The sequences correspond to SEQ ID NOs: 509, 510, 512, 520, 530, 478 from top to bottom, respectively.

[00166] **Figure 156** shows that ABE editing efficiencies improve with iterative rounds of evolution. The top panel shows representative A to G % editing at targeted genetic locus in

Hek293T cells using evolved/engineered ABE construct. The sequence corresponds to SEQ ID NO: 561. The bottom panel shows that iterative rounds of evolution and engineering improve ABE. ABE constructs are indicated by their pNMG reference numbers as shown in Table 4. The '508' target sequence corresponds to SEQ ID NO: 520.

[00167] **Figure 157** shows HTS results of core 6 genomic sites from the 10 "Best" ABE. The results indicate that different editors have different local sequence preference (bottom panel). The graph shows the A to G percent editing at 6 different genetic loci. ABE constructs are indicated by their pNMG reference numbers as shown in Table 4. The sequences correspond to SEQ ID NOs: 509, 510, 512, 520, 530, 478 from top to bottom, respectively.

[00168] **Figure 158** shows transfection of functioning "top 10" ABEs at all genomic sites covering every combination of NAN sequence. The data represents n=1. The sequences correspond to SEQ ID NOs: 489, 490, 493, 497, 503, 504, 507, 508, 511, and 513 from top to bottom, respectively.

[00169] **Figure 159** shows ABE window experiments (A's at odd positions) for identifying which A's are edited. ABEs pNMG-477, pNMG-586, pNMG-588, BE3 and untreated control are shown. The sequence for editing is shown at the top. The sequence corresponds to SEQ ID NO: 562.

[00170] **Figure 160** shows ABE window experiment (A's at even positions) for identifying which A's are edited. ABEs pNMG-477, pNMG-586, pNMG-588, BE3 and untreated control are shown. The sequence for editing is shown at the top. The sequence corresponds to SEQ ID NO: 563.

[00171] **Figure 161** shows additional ABE window experiments for identifying which A's are edited. ABEs pNMG-586, pNMG-560, and untreated control are shown. The sequence for editing is shown at the top. The sequences correspond to SEQ ID NOs: 544 and 541 from top to bottom, respectively.

[00172] **Figure 162** shows additional ABE window experiments for identifying which A's are edited. ABEs pNMG-576, pNMG-586, and untreated control are shown. The sequence for editing is shown at the top. The sequence corresponds to SEQ ID NO: 564.

[00173] **Figure 163** shows evolution # 7 an attempt to edit a multi-A site. The evolution selection design was to target 2 point mutations in the same gene using two separate gRNAs: 5'-TTCATT**A(7)**ACTGTGGCCGGCT-3'(SEQ ID NO: 565) and 5'-ATCTT**A(6)**TTCGATCATGCGAA-3' (SEQ ID NO: 566) in order to make a D208N reversion mutation in Kan and to revert a stop codon to a Q.

[00174] **Figure 164** shows evolution #7 mutations which were evolved to trget As within a multi A site, meaning that they are flanked on one or both sides by an A. The identity of mutations, relative to SEQ ID NO: 1 are shown.

[00175] **Figure 165** shows schematics of ecTadA identifying residues R152 and P48.

[00176] **Figure 166** shows MiSeq results of ABE editing on disease relevant mutations in alternative cell lines. Nucleofection with Lonza kit was used with 3 different nucleofection solutions x 16 different electroporation conditions (48 total conditions/cell line). The sequences correspond to SEQ ID NOs: 522-524 from top to bottom, respectively.

[00177] **Figure 167** shows results for A to G editing at multiple positions for various constructs. ABE constructs are indicated by their pNMG reference numbers as shown in Table 4. In the top panel the sequences correspond to SEQ ID NOs: 469-471, 567, 475, and 474 from top to bottom, respectively. In the bottom panel the sequences correspond to SEQ ID NOs: 469 (pNMG-466), 470 (pNMG-467), 471 (pNMG-469), 567 (pNMG-472), and 474 (pNMG-509) from top to bottom, respectively.

[00178] **Figure 168** shows editing results for various constructs using ABEs with different linkers. ABE constructs are indicated by their pNMG reference numbers as shown in Table 4. A schematic of the new linker ABE is also shown. The sequences correspond to SEQ ID NOs: 469 (pNMG-466), 568 (pNMG-468), 471 (pNMG-469), 567 (pNMG-472), 574 (pNGM-509), and 569) (pNMG-539) from top to bottom, respectively.

[00179] **Figure 169** shows the 4th round evolution. Evolution was done with a monomer construct and endogenous TadA complements TadA-dCas9 fusion.

[00180] **Figure 170** shows 4th round evolution results. The sequences correspond to SEQ ID NOs: 7, 368, 363, 364, 369, and 370 from top to bottom, respectively.

[00181] **Figure 171** shows evolution round #5. The plasmid and experimental outline are shown (top panel). The graph illustrates survival on chlor vs. spectinomycin "TAG" vs. "GAT." The chlor target sequence is 5'-TACGGCGT**A**GTGCACCTGGA-3' (SEQ ID NO: 441) and the spect target sequence is 5'-CAATG**A**TGACTTCTACAGCG-3'(SEQ ID NO: 444).

[00182] **Figure 172** shows editing results at the chlor and spect sites. Constructs identified from evolution #4 (site saturated/NNK library) appear edit more efficiently on the spect site rather than on the chor site. ABE constructs are indicated by their pNMG reference numbers as shown in Table 4.

[00183] **Figure 173** shows 5th round evolution (part a). The sequence corresponds to SEQ ID NO: 570.

[00184] **Figure 174** shows 5th round heterodimer (in trans) results. Round #5a identified mutations improved both editing efficiencies and broadened substrate specificity. The sequences correspond to SEQ ID NOs: 7, 368, and 363, 364, 369, and 370 from top to bottom, respectively.

[00185] **Figure 175** shows 5th round heterodimer (in cis) results. Round #5a identified mutations improved both editing efficiencies and broadened substrate specificity, but the cis results gave higher editing efficiencies. ABE constructs are indicated by their pNMG reference numbers as shown in Table 4. The sequences correspond to SEQ ID NOs: 7, 571, 465, 368, 363, 466, 364, 369, 572, and 370 from top to bottom, respectively.

[00186] **Figure 176** shows editing results of various constructs for evolution 5.

[00187] **Figure 177** shows editing results of various constructs for evolution 5.

[00188] **Figure 178** shows gRNAs for ABE. 5a constructs are characterized on all 16 NAN sequences A at position 5 in protospacer (left panel). The sequences correspond to SEQ ID NOs: 573-578 from top to bottom, respectively. Additional sequences starting with a "G" in order to minimize variations in yield gRNA synthesis are proposed (right panel). The sequences correspond to SEQ ID NOs: 579-588 from top to bottom, respectively.

[00189] **Figure 179** shows % A to G editing of A_5 using sgRNA 299 as indicated in Table 8 and the ABE constructs, which are indicated by their pNMG reference numbers as shown in Table 4. The sequence corresponds to SEQ ID NO: 478.

[00190] **Figure 180** shows % A to G editing of A_5 using sgRNA 469 as indicated in Table 8 and the ABE constructs, which are indicated by their pNMG reference numbers as shown in Table 4. The sequence corresponds to SEQ ID NO: 509.

[00191] **Figure 181** shows % A to G editing of A_5 using sgRNA 470 as indicated in Table 8 and the ABE constructs, which are indicated by their pNMG reference numbers as shown in Table 4. The sequence corresponds to SEQ ID NO: 510.

[00192] **Figure 182** shows % A to G editing of A_5 using sgRNA 472 as indicated in Table 8 and the ABE constructs, which are indicated by their pNMG reference numbers as shown in Table 4. The sequence corresponds to SEQ ID NO: 512.

[00193] **Figure 183** shows % A to G editing of A_5 using sgRNA 508 as indicated in Table 8 and the ABE constructs, which are indicated by their pNMG reference numbers as shown in Table 4. The sequence corresponds to SEQ ID NO: 520.

[00194] **Figure 184** shows % A to G editing of A_7 using sgRNA 536 as indicated in Table 8 and the ABE constructs, which are indicated by their pNMG reference numbers as shown in Table 4. The sequence corresponds to SEQ ID NO: 530.

[00195] **Figure 185** shows the % of A to G editing of the highlighted A (A_5) using sgRNA: 310, sgRNA: 311, sgRNA: 314, sgRNA: 318, sgRNA: 463, and sgRNA: 464 for each of the indicated base editors, which are indicated by their pNMG reference numbers as shown in Table 4. The sequences correspond to SEQ ID NOs: 489, 490, 493, 497, 503 and 504 from left to right and top to bottom, respectively.

[00196] **Figure 186** shows the % of A to G editing of the highlighted A (A_5) using sgRNA: 466, sgRNA: 467, sgRNA: 468, sgRNA: 471, sgRNA: 501, and sgRNA: 601 for each of the indicated base editors, which are indicated by their pNMG reference numbers as shown in Table 4. The sequences correspond to SEQ ID NOs: 506, 507, 508, 511, 513, and 535 from left to right and top to bottom, respectively.

DEFINITIONS

[00197] As used herein and in the claims, the singular forms "a," "an," and "the" include the singular and the plural unless the context clearly indicates otherwise. Thus, for example, a reference to "an agent" includes a single agent and a plurality of such agents.

[00198] The term "deaminase" or "deaminase domain" refers to a protein or enzyme that catalyzes a deamination reaction. In some embodiments, the deaminase is an adenosine deaminase, which catalyzes the hydrolytic deamination of adenine or adenosine. In some embodiments, the deaminase or deaminase domain is an adenosine deaminase, catalyzing the hydrolytic deamination of adenosine or deoxyadenosine to inosine or deoxyinosine, respectively. In some embodiments, the adenosine deaminase catalyzes the hydrolytic deamination of adenine or adenosine in deoxyribonucleic acid (DNA). The adenosine deaminases (*e.g.* engineered adenosine deaminases, evolved adenosine deaminases) provided herein may be from any organism, such as a bacterium. In some embodiments, the deaminase or deaminase domain is a variant of a naturally-occurring deaminase from an organism. In some embodiments, the deaminase or deaminase domain does not occur in nature. For example, in some embodiments, the deaminase or deaminase domain is 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%, at least 96%, at least 97%, at least 98%, at least 99%, or at least 99.5% identical to a naturally-occurring deaminase. In some embodiments, the adenosine deaminase is from a bacterium, such as, *E.coli*, *S. aureus, S. typhi, S. putrefaciens, H. influenzae,* or *C. crescentus*. In some embodiments, the adenosine deaminase is a TadA deaminase. In some embodiments, the TadA deaminase is an *E. coli* TadA deaminase (ecTadA). In some embodiments, the TadA deaminase is a truncated *E. coli* TadA

deaminase. For example, the truncated ecTadA may be missing one or more N-terminal amino acids relative to a full-length ecTadA. In some embodiments, the truncated ecTadA may be missing 1, 2, 3, 4, 5 ,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, or 20 N-terminal amino acid residues relative to the full length ecTadA. In some embodiments, the truncated ecTadA may be missing 1, 2, 3, 4, 5 ,6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 6, 17, 18, 19, or 20 C-terminal amino acid residues relative to the full length ecTadA. In some embodiments, the ecTadA deaminase does not comprise an N-terminal methionine

[00199] In some embodiments, the TadA deaminase is an N-terminal truncated TadA. In certain embodiments, the adenosine deaminase comprises the amino acid sequence: MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPT AHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKT GAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD (SEQ ID NO: 1).

[00200] In some embodiments the TadA deaminase is a full-length *E. coli* TadA deaminase. For example, in certain embodiments, the adenosine deaminase comprises the amino acid sequence:
MRRAFITGVFFLSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIG RVVFGARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEI KAQKKAQSSTD (SEQ ID NO: 84)

[00201] It should be appreciated, however, that additional adenosine deaminases useful in the present application would be apparent to the skilled artisan and are within the scope of this disclosure. For example, the adenosine deaminase may be a homolog of an ADAT. Exemplary ADAT homologs include, without limitation:

Staphylococcus aureus TadA:
MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAH AEHIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCS GSLMNLLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 8)

Bacillus subtilis TadA:
MTQDELYMKEAIKEAKKAEEKGEVPIGAVLVINGEIIARAHNLRETEQRSIAHAEML VID

EACKALGTWRLEGATLYVTLEPCPMCAGAVVLSRVEKVVFGAFDPKGGCSGTLMN LLQEERFNHQAEVVSGVLEEECGGMLSAFFRELRKKKKAARKNLSE (SEQ ID NO: 9)

Salmonella typhimurium (*S. typhimurium*) TadA:
MPPAFITGVTSLSDVELDHEYWMRHALTLAKRAWDEREVPVGAVLVHNHRVIGEG WNRPIGRHDPTAHAEIMALRQGGLVLQNYRLLDTTLYVTLEPCVMCAGAMVHSRIG RVVFGARDAKTGAAGSLIDVLHHPGMNHRVEIIEGVLRDECATLLSDFFRMRRQEIK ALKKADRAEGAGPAV (SEQ ID NO: 371)

Shewanella putrefaciens (*S. putrefaciens*) TadA:
MDEYWMQVAMQMAEKAEAAGEVPVGAVLVKDGQQIATGYNLSISQHDPTAHAEI LCLRSAGKKLENYRLLDATLYITLEPCAMCAGAMVHSRIARVVYGARDEKTGAAGT VVNLLQHPAFNHQVEVTSGVLAEACSAQLSRFFKRRRDEKKALKLAQRAQQGIE (SEQ ID NO: 372)

Haemophilus influenzae F3031 (*H. influenzae*) TadA:
MDAAKVRSEFDEKMMRYALELADKAEALGEIPVGAVLVDDARNIIGEGWNLSIVQS DPTAHAEIIALRNGAKNIQNYRLLNSTLYVTLEPCTMCAGAILHSRIKRLVFGASDYK TGAIGSRFHFFDDYKMNHTLEITSGVLAEECSQKLSTFFQKRREEKKIEKALLKSLSD K (SEQ ID NO: 373)

Caulobacter crescentus (C. crescentus) TadA:
MRTDESEDQDHRMMRLALDAARAAAEAGETPVGAVILDPSTGEVIATAGNGPIAAH DPTAHAEIAAMRAAAAKLGNYRLTDLTLVVTLEPCAMCAGAISHARIGRVVFGADD PKGGAVVHGPKFFAQPTCHWRPEVTGGVLADESADLLRGFFRARRKAKI (SEQ ID NO: 374)

Geobacter sulfurreducens (*G. sulfurreducens*) TadA:
MSSLKKTPIRDDAYWMGKAIREAAKAAARDEVPIGAVIVRDGAVIGRGHNLREGSN DPSAHAEMIAIRQAARRSANWRLTGATLYVTLEPCLMCMGAIILARLERVVFGCYDP KGGAAGSLYDLSADPRLNHQVRLSPGVCQEECGTMLSDFFRDLRRRKKAKATPALF IDERKVPPEP (SEQ ID NO: 375)

[00202] The term "base editor (BE)," or "nucleobase editor (NBE)" refers to an agent comprising a polypeptide that is capable of making a modification to a base (e.g., A, T, C, G, or U) within a nucleic acid sequence (*e.g.*, DNA or RNA). In some embodiments, the base editor is capable of deaminating a base within a nucleic acid. In some embodiments, the base editor is capable of deaminating a base within a DNA molecule. In some embodiments, the base editor is capable of deaminating an adenine (A) in DNA. In some embodiments, the base editor is a fusion protein comprising a nucleic acid programmable DNA binding protein (napDNAbp) fused to an adenosine deaminase. In some embodiments, the base editor is a Cas9 protein fused to an adenosine deaminase. In some embodiments, the base editor is a Cas9 nickase (nCas9) fused to an adenosine deaminase. In some embodiments, the base editor is a nuclease-inactive Cas9 (dCas9) fused to an adenosine deaminase. In some embodiments, the base editor is fused to an inhibitor of base excision repair, for example, a UGI domain, or a dISN domain. In some embodiments, the fusion protein comprises a Cas9 nickase fused to a deaminase and an inhibitor of base excision repair, such as a UGI or dISN domain. In some embodiments, the dCas9 domain of the fusion protein comprises a D10A and a H840A mutation of SEQ ID NO: 52, or a corresponding mutation in any of SEQ ID NOs: 108-357, which inactivates the nuclease activity of the Cas9 protein. In some embodiments, the fusion protein comprises a D10A mutation and comprises a histidine at residue 840 of SEQ ID NO: 52, or a corresponding mutation in any of SEQ ID NOs: 108-357, which renders Cas9 capable of cleaving only one strand of a nucleic acid duplex. An example of a Cas9 nickase is shown in SEQ ID NO: 35.

[00203] The term "linker," as used herein, refers to a bond (*e.g.*, covalent bond), chemical group, or a molecule linking two molecules or moieties, *e.g.*, two domains of a fusion protein, such as, for example, a nuclease-inactive Cas9 domain and a nucleic acid-editing domain (*e.g.*, an adenosine deaminase). In some embodiments, a linker joins a gRNA binding domain of an RNA-programmable nuclease, including a Cas9 nuclease domain, and the catalytic domain of a nucleic-acid editing protein. In some embodiments, a linker joins a dCas9 and a nucleic-acid editing protein. 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. In some embodiments, a linker comprises the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 10), which may also be referred to as the XTEN linker. In some embodiments, a linker comprises the amino acid sequence SGGS (SEQ ID NO: 37). In some embodiments, a linker comprises $(SGGS)_n$ (SEQ ID NO: 37), $(GGGS)_n$ (SEQ ID NO: 38), $(GGGGS)_n$ (SEQ ID NO: 39), $(G)_n$, $(EAAAK)_n$ (SEQ ID NO: 40), $(GGS)_n$, SGSETPGTSESATPES (SEQ ID NO: 10), or $(XP)_n$ motif, or a combination of any of these, wherein n is independently an integer between 1 and 30, and wherein X is any amino acid. In some embodiments, n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15.

[00204] 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* (4^{th} ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (2012)).

[00205] 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 inosine base excision repair. Exemplary inhibitors of base 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 a catalytically inactive EndoV or a catalytically inactive hAAG.

[00206] The term "uracil glycosylase inhibitor" or "UGI," 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: 3. 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: 3. 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: 3. In some embodiments,

a UGI comprises an amino acid sequence homologous to the amino acid sequence set forth in SEQ ID NO: 3, or an amino acid sequence homologous to a fragment of the amino acid sequence set forth in SEQ ID NO: 3. 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: 3. 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: 3. In some embodiments, the UGI comprises the following amino acid sequence: >sp|P14739|UNGI_BPPB2 Uracil-DNA glycosylase inhibitor MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLT SDAPEYKPWALVIQDSNGENKIKML (SEQ ID NO: 3).

[00207] The term "catalytically inactive inosine-specific nuclease," or "dead inosine-specific nuclease (dISN)," as used herein, refers to a protein that is capable of inhibiting an inosine-specific nuclease. Without wishing to be bound by any particular theory, catalytically inactive inosine glycosylases (*e.g.,* alkyl adenine glycosylase [AAG]) will bind inosine, but will not create an abasic site or remove the inosine, thereby sterically blocking the newly-formed inosine moiety from DNA damage/repair mechanisms. In some embodiments, the catalytically inactive inosine-specific nuclease may be capable of binding an inosine in a nucleic acid but does not cleave the nucleic acid. Exemplary catalytically inactive inosine-specific nucleases include, without limitation, catalytically inactive alkyl adenosine glycosylase (AAG nuclease), for example, from a human, and catalytically inactive endonuclease V (EndoV nuclease), for example, from *E. coli*. In some embodiments, the catalytically inactive AAG nuclease comprises an E125Q mutation as shown in SEQ ID NO: 32, or a corresponding mutation in another AAG nuclease. In some embodiments, the catalytically inactive AAG nuclease comprises the amino acid sequence set forth in SEQ ID NO: 32. In some embodiments, the catalytically inactive EndoV nuclease comprises an D35A mutation as shown in SEQ ID NO 32, or a corresponding mutation in another EndoV

nuclease. In some embodiments, the catalytically inactive EndoV nuclease comprises the amino acid sequence set forth in SEQ ID NO: 33. It should be appreciated that other catalytically inactive inosine-specific nucleases (dISNs) would be apparent to the skilled artisan and are within the scope of this disclosure.

Truncated AAG (*H. sapiens*) nuclease (E125Q); mutated residue underlined in bold.

[00208] KGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELRGRIVET**Q**AY LGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDGACVLLRA LEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRDLAQDE AVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRVAEQ DTQA (SEQ ID NO: 32)

EndoV nuclease (D35A); mutated residue underlined in bold.

[00209] DLASLRAQQIELASSVIREDRLDKDPPDLIAGA**A**VGFEQGGEVTRAAM VLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVFVDGHG ISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQLAWV WRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPAFVR YTANQP (SEQ ID NO: 33)

[00210] 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 asWO/2001/038547 on May 31, 2001, the contents of which are incorporated herein by reference for their disclosure of exemplary nuclear localization sequences. In some embodiments, a NLS comprises the amino acid sequence PKKKRKV (SEQ ID NO: 4) or MDSLLMNRRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 5).

[00211] The term "nucleic acid programmable DNA binding protein" or "napDNAbp" refers to a protein that associates with a nucleic acid (*e.g.,* DNA or RNA), such as a guide nuclic acid, that guides the napDNAbp to a specific nucleic acid sequence. For example, a Cas9 protein can associate with a guide RNA that guides the Cas9 protein to a specific DNA sequence that has complementary to the guide RNA. In some embodiments, the napDNAbp is a class 2 microbial CRISPR-Cas effector. In some embodiments, the napDNAbp is a Cas9

domain, for example a nuclease active Cas9, a Cas9 nickase (nCas9), or a nuclease inactive Cas9 (dCas9). Examples of nucleic acid programmable DNA binding proteins include, without limitation, Cas9 (*e.g.*, dCas9 and nCas9), CasX, CasY, Cpf1, C2c1, C2c2, C2C3, and Argonaute. It should be appreciated, however, that nucleic acid programmable DNAbinding proteins also include nucleic acid programmable proteins that bind RNA. For example, the napDNAbp may be associated with a nucleic acid that guides the napDNAbp to an RNA. Other nucleic acid programmable DNA binding proteins are also within the scope of this disclosure, though they may not be specifically listed in this disclosure.

[00212] The term "Cas9" or "Cas9 domain" refers to an RNA-guided nuclease comprising a Cas9 protein, or a fragment thereof (*e.g.*, a protein comprising an active, inactive, or partially active DNA cleavage domain of Cas9, and/or the gRNA binding domain of Cas9). 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 protein. The tracrRNA serves as a guide for ribonuclease 3-aided processing of pre-crRNA. Subsequently, Cas9/crRNA/tracrRNA endonucleolytically cleaves 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 so as 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 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. 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 has an inactive (*e.g.*, an inactivated) DNA cleavage domain, that is, the Cas9 is a nickase.

[00213] A nuclease-inactivated Cas9 protein may interchangeably be referred to as a "dCas9" protein (for nuclease-"dead" Cas9). Methods for generating a Cas9 protein (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, or at least about

99.9% identical to wild type Cas9. 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. In some embodiments, the Cas9 variant comprises a fragment of 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.

[00214] In some embodiments, the 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 1300 amino acids in length. In some embodiments, wild type Cas9 corresponds to Cas9 from *Streptococcus pyogenes* (NCBI Reference Sequence: NC_017053.1, SEQ ID NO: 47 (nucleotide); SEQ ID NO: 48 (amino acid)).

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGG
GCGGTGATCACTGATGATTATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAA
ATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTTTATTTGGCAG
TGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATAC
ACGTCGGAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCG
AAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAG
ACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTA
TCATGAGAAATATCCAACTATCTATCATCTGCGAAAAAAATTGGCAGATTCTACT
GATAAAGCGGATTTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAAGTTTC
GTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGATGTGGACAA
ACTATTTATCCAGTTGGTACAAATCTACAATCAATTATTTGAAGAAAACCCTATT
AACGCAAGTAGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCA
AGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAGAAATGGCTTG
TTTGGGAATCTCATTGCTTTGTCATTGGGATTGACCCCTAATTTTAAATCAAATTT
TGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGAT

TTAGATAATTTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAG
CTAAGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAGAGTAAATAGTGA
AATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAGCGCTACGATGAACATCAT
CAAGACTTGACTCTTTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATA
AAGAAATCTTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGG
AGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGAT
GGTACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAA
CGGACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGTGAGCTGCATG
CTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAATCGTGAGAA
GATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTG
GCAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATG
GAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATTGAACGC
ATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGT
TTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTA
CTGAGGGAATGCGAAAACCAGCATTTCTTTCAGGTGAACAGAAGAAAGCCATTG
TTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAG
ATTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGA
TAGATTTAATGCTTCATTAGGCGCCTACCATGATTTGCTAAAAATTATTAAAGAT
AAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAA
CATTGACCTTATTTGAAGATAGGGGGATGATTGAGGAAAGACTTAAAACATATG
CTCACCTCTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGG
TTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGATAAGCAATCTGGC
AAAACAATATTAGATTTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTTATGC
AGCTGATCCATGATGATAGTTTGACATTTAAAGAAGATATTCAAAAAGCACAGG
TGTCTGGACAAGGCCATAGTTTACATGAACAGATTGCTAACTTAGCTGGCAGTCC
TGCTATTAAAAAAGGTATTTTACAGACTGTAAAAATTGTTGATGAACTGGTCAAA
GTAATGGGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAATCAG
ACAACTCAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGAAGA
AGGTATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAATAC
TCAATTGCAAAATGAAAAGCTCTATCTCTATTATCTACAAAATGGAAGAGACATG
TATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGATCACA
TTGTTCCACAAAGTTTCATTAAAGACGATTCAATAGACAATAAGGTACTAACGCG
TTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGTCAA
AAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCAACG

TAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAACTTGATAAA
GCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCACTAAGCATGTGG
CACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAACTTAT
TCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTCCGAAAA
GATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCATGATG
CGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAACTTGA
ATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGATTGCT
AAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTAATA
TCATGAACTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTCGCAAAC
GCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATAAAGGGC
GAGATTTTGCCACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTCAA
GAAAACAGAAGTACAGACAGGCGGATTCTCCAAGGAGTCAATTTTACCAAAAAG
AAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCAAAAAAATATGG
TGGTTTTGATAGTCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGGAA
AAAGGGAAATCGAAGAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAATT
ATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTGACTTTTTAGAAGCTAAAGGAT
ATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTTGA
GTTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAAGG
AAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGTCAT
TATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTGTG
GAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTCTA
AGCGTGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAACAA
ACATAGAGACAAACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTTAC
GTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTGATC
GTAAACGATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCAATC
CATCACTGGTCTTTATGAAACACGCATTGATTTGAGTCAGCTAGGAGGTGACTGA
(SEQ ID NO:47)

MDKKYSIGLDIGTNSVGWAVITDDYKVPSKKFKVLGNTDRHSIKKNLIGALLFGSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLADSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQIYNQLFEENPINASRVDAKAILSARLSKSRRLENLIAQLPG
EKRNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYAD
LFLAAKNLSDAILLSDILRVNSEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEK

YKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV
YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGAYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDRGMIEER
LKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANR
NFMQLIHDDSLTFKEDIQKAQVSGQGHSLHEQIANLAGSPAIKKGILQTVKIVDELVK
VMGHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQ
NEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFIKDDSIDNKVLTRSDKNR
GKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKRQ
LVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKVREI
NNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKAT
AKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSMPQ
VNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELE
NGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHK
HYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:48)
(single underline: HNH domain; double underline: RuvC domain)

[00215] In some embodiments, wild type Cas9 corresponds to, or comprises SEQ ID NO:49 (nucleotide) and/or SEQ ID NO: 50 (amino acid):

ATGGATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCGTTGGATGGG
CTGTCATAACCGATGAATACAAAGTACCTTCAAAGAAATTTAAGGTGTTGGGGA
ACACAGACCGTCATTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTCGATAG
TGGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATAC
ACGTCGCAAGAACCGAATATGTTACTTACAAGAAATTTTTAGCAATGAGATGGCC
AAAGTTGACGATTCTTTCTTTCACCGTTTGGAAGAGTCCTTCCTTGTCGAAGAGG
ACAAGAAACATGAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCAT
ATCATGAAAAGTACCCAACGATTTATCACCTCAGAAAAAAGCTAGTTGACTCAA
CTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATAAAGTT
CCGTGGGCACTTTCTCATTGAGGGTGATCTAAATCCGGACAACTCGGATGTCGAC
AAACTGTTCATCCAGTTAGTACAAACCTATAATCAGTTGTTTGAAGAGAACCCTA
TAAATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCTCTAAATC
CCGACGGCTAGAAAACCTGATCGCACAATTACCCGGAGAGAAGAAAAATGGGTT

GTTCGGTAACCTTATAGCGCTCTCACTAGGCCTGACACCAAATTTTAAGTCGAAC
TTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTACGATGAC
GATCTCGACAATCTACTGGCACAAATTGGAGATCAGTATGCGGACTTATTTTTGG
CTGCCAAAAACCTTAGCGATGCAATCCTCCTATCTGACATACTGAGAGTTAATAC
TGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTACGATGAACAT
CACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTCAGCAACTGCCTGAGAAAT
ATAAGGAAATATTCTTTGATCAGTCGAAAAACGGGTACGCAGGTTATATTGACG
GCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATATTAGAGAAGA
TGGATGGGACGGAAGAGTTGCTTGTAAAACTCAATCGCGAAGATCTACTGCGAA
AGCAGCGGACTTTCGACAACGGTAGCATTCCACATCAAATCCACTTAGGCGAATT
GCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCCTCAAAGACAATCGT
GAAAAGATTGAGAAAATCCTAACCTTTCGCATACCTTACTATGTGGGACCCCTGG
CCCGAGGGAACTCTCGGTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTA
CTCCATGGAATTTTGAGGAAGTTGTCGATAAAGGTGCGTCAGCTCAATCGTTCAT
CGAGAGGATGACCAACTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAA
GCACAGTTTACTTTACGAGTATTTCACAGTGTACAATGAACTCACGAAAGTTAAG
TATGTCACTGAGGGCATGCGTAAACCCGCCTTTCTAAGCGGAGAACAGAAGAAA
GCAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTG
AAAGAGGACTACTTTAAGAAAATTGAATGCTTCGATTCTGTCGAGATCTCCGGGG
TAGAAGATCGATTTAATGCGTCACTTGGTACGTATCATGACCTCCTAAAGATAAT
TAAAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATAT
AGTGTTGACTCTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAA
ACATACGCTCACCTGTTCGACGATAAGGTTATGAAACAGTTAAAGAGGCGTCGCT
ATACGGGCTGGGGACGATTGTCGCGGAAACTTATCAACGGGATAAGAGACAAGC
AAAGTGGTAAAACTATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAA
CTTTATGCAGCTGATCCATGATGACTCTTTAACCTTCAAAGAGGATATACAAAAG
GCACAGGTTTCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTG
GTTCGCCAGCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGC
TAGTTAAGGTCATGGGACGTCACAAACCGGAAAACATTGTAATCGAGATGGCAC
GCGAAAATCAAACGACTCAGAAGGGGCAAAAAAACAGTCGAGAGCGGATGAAG
AGAATAGAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCT
GTGGAAAATACCCAATTGCAGAACGAGAAACTTTACCTCTATTACCTACAAAATG
GAAGGGACATGTATGTTGATCAGGAACTGGACATAAACCGTTTATCTGATTACGA
CGTCGATCACATTGTACCCCAATCCTTTTTGAAGGACGATTCAATCGACAATAAA

GTGCTTACACGCTCGGATAAGAACCGAGGGAAAAGTGACAATGTTCCAAGCGAG
GAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAAACTG
ATAACGCAAAGAAAGTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCT
GAACTTGACAAGGCCGGATTTATTAAACGTCAGCTCGTGGAAACCCGCCAAATC
ACAAAGCATGTTGCACAGATACTAGATTCCCGAATGAATACGAAATACGACGAG
AACGATAAGCTGATTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAAATTGGTG
TCGGACTTCAGAAAGGATTTTCAATTCTATAAAGTTAGGGAGATAAATAACTACC
ACCATGCGCACGACGCTTATCTTAATGCCGTCGTAGGGACCGCACTCATTAAGAA
ATACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTC
CGTAAGATGATCGCGAAAAGCGAACAGGAGATAGGCAAGGCTACAGCCAAATA
CTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAACG
GAGAGATACGCAAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCG
TATGGGATAAGGGCCGGGACTTCGCGACGGTGAGAAAAGTTTTGTCCATGCCCC
AAGTCAACATAGTAAAGAAAACTGAGGTGCAGACCGGAGGGTTTTCAAAGGAAT
CGATTCTTCCAAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGG
ACCCGAAAAAGTACGGTGGCTTCGATAGCCCTACAGTTGCCTATTCTGTCCTAGT
AGTGGCAAAAGTTGAGAAGGGAAAATCCAAGAAACTGAAGTCAGTCAAAGAAT
TATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAAAGAACCCCATCGACTT
CCTTGAGGCGAAAGGTTACAAGGAAGTAAAAAAGGATCTCATAATTAAACTACC
AAAGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGCGC
CGGAGAGCTTCAAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTT
CCTGTATTTAGCGTCCCATTACGAGAAGTTGAAAGGTTCACCTGAAGATAACGAA
CAGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATAGAGC
AAATTTCGGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGACAAAGT
ATTAAGCGCATACAACAAGCACAGGGATAAACCCATACGTGAGCAGGCGGAAA
ATATTATCCATTTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTCAAGTAT
TTTGACACAACGATAGATCGCAAACGATACACTTCTACCAAGGAGGTGCTAGAC
GCGACACTGATTCACCAATCCATCACGGGATTATATGAAACTCGGATAGATTTGT
CACAGCTTGGGGGTGACGGATCCCCCAAGAAGAAGAGGAAAGTCTCGAGCGACT
ACAAAGACCATGACGGTGATTATAAAGATCATGACATCGATTACAAGGATGACG
ATGACAAGGCTGCAGGA (SEQ ID NO: 49)

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE

RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR
TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV
YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:
50) (single underline: HNH domain; double underline: RuvC domain)

[00216] In some embodiments, wild type Cas9 corresponds to Cas9 from *Streptococcus pyogenes* (NCBI Reference Sequence: NC_002737.2, SEQ ID NO: 51 (nucleotide); and Uniport Reference Sequence: Q99ZW2, SEQ ID NO: 52 (amino acid).

ATGGATAAGAAATACTCAATAGGCTTAGATATCGGCACAAATAGCGTCGGATGG
GCGGTGATCACTGATGAATATAAGGTTCCGTCTAAAAAGTTCAAGGTTCTGGGAA
ATACAGACCGCCACAGTATCAAAAAAAATCTTATAGGGGCTCTTTTATTTGACAG
TGGAGAGACAGCGGAAGCGACTCGTCTCAAACGGACAGCTCGTAGAAGGTATAC
ACGTCGGAAGAATCGTATTTGTTATCTACAGGAGATTTTTTCAAATGAGATGGCG
AAAGTAGATGATAGTTTCTTTCATCGACTTGAAGAGTCTTTTTTGGTGGAAGAAG
ACAAGAAGCATGAACGTCATCCTATTTTTGGAAATATAGTAGATGAAGTTGCTTA
TCATGAGAAATATCCAACTATCTATCATCTGCGAAAAAAATTGGTAGATTCTACT

GATAAAGCGGATTTGCGCTTAATCTATTTGGCCTTAGCGCATATGATTAAGTTTC
GTGGTCATTTTTTGATTGAGGGAGATTTAAATCCTGATAATAGTGATGTGGACAA
ACTATTTATCCAGTTGGTACAAACCTACAATCAATTATTTGAAGAAAACCCTATT
AACGCAAGTGGAGTAGATGCTAAAGCGATTCTTTCTGCACGATTGAGTAAATCA
AGACGATTAGAAAATCTCATTGCTCAGCTCCCCGGTGAGAAGAAAAATGGCTTA
TTTGGGAATCTCATTGCTTTGTCATTGGGTTTGACCCCTAATTTTAAATCAAATTT
TGATTTGGCAGAAGATGCTAAATTACAGCTTTCAAAAGATACTTACGATGATGAT
TTAGATAATTTATTGGCGCAAATTGGAGATCAATATGCTGATTTGTTTTTGGCAG
CTAAGAATTTATCAGATGCTATTTTACTTTCAGATATCCTAAGAGTAAATACTGA
AATAACTAAGGCTCCCCTATCAGCTTCAATGATTAAACGCTACGATGAACATCAT
CAAGACTTGACTCTTTTAAAAGCTTTAGTTCGACAACAACTTCCAGAAAAGTATA
AAGAAATCTTTTTTGATCAATCAAAAAACGGATATGCAGGTTATATTGATGGGGG
AGCTAGCCAAGAAGAATTTTATAAATTTATCAAACCAATTTTAGAAAAAATGGAT
GGTACTGAGGAATTATTGGTGAAACTAAATCGTGAAGATTTGCTGCGCAAGCAA
CGGACCTTTGACAACGGCTCTATTCCCCATCAAATTCACTTGGGTGAGCTGCATG
CTATTTTGAGAAGACAAGAAGACTTTTATCCATTTTTAAAAGACAATCGTGAGAA
GATTGAAAAAATCTTGACTTTTCGAATTCCTTATTATGTTGGTCCATTGGCGCGTG
GCAATAGTCGTTTTGCATGGATGACTCGGAAGTCTGAAGAAACAATTACCCCATG
GAATTTTGAAGAAGTTGTCGATAAAGGTGCTTCAGCTCAATCATTTATTGAACGC
ATGACAAACTTTGATAAAAATCTTCCAAATGAAAAAGTACTACCAAAACATAGT
TTGCTTTATGAGTATTTTACGGTTTATAACGAATTGACAAAGGTCAAATATGTTA
CTGAAGGAATGCGAAAACCAGCATTTCTTTCAGGTGAACAGAAGAAAGCCATTG
TTGATTTACTCTTCAAAACAAATCGAAAAGTAACCGTTAAGCAATTAAAAGAAG
ATTATTTCAAAAAAATAGAATGTTTTGATAGTGTTGAAATTTCAGGAGTTGAAGA
TAGATTTAATGCTTCATTAGGTACCTACCATGATTTGCTAAAAATTATTAAAGAT
AAAGATTTTTTGGATAATGAAGAAAATGAAGATATCTTAGAGGATATTGTTTTAA
CATTGACCTTATTTGAAGATAGGGAGATGATTGAGGAAAGACTTAAAACATATG
CTCACCTCTTTGATGATAAGGTGATGAAACAGCTTAAACGTCGCCGTTATACTGG
TTGGGGACGTTTGTCTCGAAAATTGATTAATGGTATTAGGGATAAGCAATCTGGC
AAAACAATATTAGATTTTTTGAAATCAGATGGTTTTGCCAATCGCAATTTTATGC
AGCTGATCCATGATGATAGTTTGACATTTAAAGAAGACATTCAAAAAGCACAAG
TGTCTGGACAAGGCGATAGTTTACATGAACATATTGCAAATTTAGCTGGTAGCCC
TGCTATTAAAAAAGGTATTTTACAGACTGTAAAAGTTGTTGATGAATTGGTCAAA
GTAATGGGGCGGCATAAGCCAGAAAATATCGTTATTGAAATGGCACGTGAAAAT

CAGACAACTCAAAAGGGCCAGAAAAATTCGCGAGAGCGTATGAAACGAATCGA
AGAAGGTATCAAAGAATTAGGAAGTCAGATTCTTAAAGAGCATCCTGTTGAAAA
TACTCAATTGCAAAATGAAAAGCTCTATCTCTATTATCTCCAAAATGGAAGAGAC
ATGTATGTGGACCAAGAATTAGATATTAATCGTTTAAGTGATTATGATGTCGATC
ACATTGTTCCACAAAGTTTCCTTAAAGACGATTCAATAGACAATAAGGTCTTAAC
GCGTTCTGATAAAAATCGTGGTAAATCGGATAACGTTCCAAGTGAAGAAGTAGT
CAAAAAGATGAAAAACTATTGGAGACAACTTCTAAACGCCAAGTTAATCACTCA
ACGTAAGTTTGATAATTTAACGAAAGCTGAACGTGGAGGTTTGAGTGAACTTGAT
AAAGCTGGTTTTATCAAACGCCAATTGGTTGAAACTCGCCAAATCACTAAGCATG
TGGCACAAATTTTGGATAGTCGCATGAATACTAAATACGATGAAAATGATAAAC
TTATTCGAGAGGTTAAAGTGATTACCTTAAAATCTAAATTAGTTTCTGACTTCCG
AAAAGATTTCCAATTCTATAAAGTACGTGAGATTAACAATTACCATCATGCCCAT
GATGCGTATCTAAATGCCGTCGTTGGAACTGCTTTGATTAAGAAATATCCAAAAC
TTGAATCGGAGTTTGTCTATGGTGATTATAAAGTTTATGATGTTCGTAAAATGATT
GCTAAGTCTGAGCAAGAAATAGGCAAAGCAACCGCAAAATATTTCTTTTACTCTA
ATATCATGAACTTCTTCAAAACAGAAATTACACTTGCAAATGGAGAGATTCGCAA
ACGCCCTCTAATCGAAACTAATGGGGAAACTGGAGAAATTGTCTGGGATAAAGG
GCGAGATTTTGCCACAGTGCGCAAAGTATTGTCCATGCCCCAAGTCAATATTGTC
AAGAAAACAGAAGTACAGACAGGCGGATTCTCCAAGGAGTCAATTTTACCAAAA
AGAAATTCGGACAAGCTTATTGCTCGTAAAAAAGACTGGGATCCAAAAAAATAT
GGTGGTTTTGATAGTCCAACGGTAGCTTATTCAGTCCTAGTGGTTGCTAAGGTGG
AAAAAGGGAAATCGAAGAAGTTAAAATCCGTTAAAGAGTTACTAGGGATCACAA
TTATGGAAAGAAGTTCCTTTGAAAAAAATCCGATTGACTTTTTAGAAGCTAAAGG
ATATAAGGAAGTTAAAAAAGACTTAATCATTAAACTACCTAAATATAGTCTTTTT
GAGTTAGAAAACGGTCGTAAACGGATGCTGGCTAGTGCCGGAGAATTACAAAAA
GGAAATGAGCTGGCTCTGCCAAGCAAATATGTGAATTTTTTATATTTAGCTAGTC
ATTATGAAAAGTTGAAGGGTAGTCCAGAAGATAACGAACAAAAACAATTGTTTG
TGGAGCAGCATAAGCATTATTTAGATGAGATTATTGAGCAAATCAGTGAATTTTC
TAAGCGTGTTATTTTAGCAGATGCCAATTTAGATAAAGTTCTTAGTGCATATAAC
AAACATAGAGACAAACCAATACGTGAACAAGCAGAAAATATTATTCATTTATTT
ACGTTGACGAATCTTGGAGCTCCCGCTGCTTTTAAATATTTTGATACAACAATTG
ATCGTAAACGATATACGTCTACAAAAGAAGTTTTAGATGCCACTCTTATCCATCA
ATCCATCACTGGTCTTTATGAAACACGCATTGATTTGAGTCAGCTAGGAGGTGAC
TGA (SEQ ID NO: 51)

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR
TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV
YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:
52) (single underline: HNH domain; double underline: RuvC domain)

[00217] In some embodiments, Cas9 refers to Cas9 from: *Corynebacterium ulcerans* (NCBI Refs: NC_015683.1, NC_017317.1); *Corynebacterium diphtheria* (NCBI Refs: NC_016782.1, NC_016786.1); *Spiroplasma syrphidicola* (NCBI Ref: NC_021284.1); *Prevotella intermedia* (NCBI Ref: NC_017861.1); *Spiroplasma taiwanense* (NCBI Ref: NC_021846.1); *Streptococcus iniae* (NCBI Ref: NC_021314.1); *Belliella baltica* (NCBI Ref: NC_018010.1); *Psychroflexus torquisI* (NCBI Ref: NC_018721.1); *Streptococcus thermophilus* (NCBI Ref: YP_820832.1), *Listeria innocua* (NCBI Ref: NP_472073.1),

Campylobacter jejuni (NCBI Ref: YP_002344900.1) or *Neisseria. meningitidis* (NCBI Ref: YP_002342100.1) or to a Cas9 from any other organism.

[00218] In some 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. For example, in some embodiments, a dCas9 domain comprises D10A and an H840A mutation of SEQ ID NO: 52 or corresponding mutations in another Cas9. In some embodiments, the dCas9 comprises the amino acid sequence of SEQ ID NO: 53

dCas9 (D10A and H840A):

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR
TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV
YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:
53) (single underline: HNH domain; double underline: RuvC domain).

[00219] In some embodiments, the Cas9 domain comprises a D10A mutation, while the residue at position 840 remains a histidine in the amino acid sequence provided in SEQ ID NO: 52, or at corresponding positions in any of the amino acid sequences provided in SEQ ID NOs: 108-357. Without wishing to be bound by any particular theory, the presence of the catalytic residue H840 maintains the activity of the Cas9 to cleave the non-edited (*e.g.*, non-deaminated) strand containing a T opposite the targeted A. Restoration of H840 (*e.g.*, from A840 of a dCas9) does not result in the cleavage of the target strand containing the A. Such Cas9 variants are able to generate a single-strand DNA break (nick) at a specific location based on the gRNA-defined target sequence, leading to repair of the non-edited strand, ultimately resulting in a T to C change on the non-edited strand. A schematic representation of this process is shown in Figure 94. Briefly, and without wishing to be bound by any particular theory, the A of a A-T base pair can be deaminated to a inosine (I) by an adenosine deaminase, *e.g.,* an engineered adenosine deaminase that deaminates an adenosine in DNA. Nicking the non-edited strand, having the T, facilitates removal of the T via mismatch repair mechanisms. A UGI domain or a catalytically inactive inosine-specific nuclease (dISN) may inhibit inosine-specific nucleases (*e.g.*, sterically) thereby preventing removal of the inosine (I).

[00220] In other embodiments, dCas9 variants having mutations other than D10A and H840A are provided, which, *e.g.*, result in nuclease inactivated Cas9 (dCas9). 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). In some embodiments, variants or homologues of dCas9 (*e.g.*, variants of SEQ ID NO: 53) 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 SEQ ID NO: 10. In some embodiments, variants of dCas9 (*e.g.*, variants of SEQ ID NO: 53) are provided having amino acid sequences which are shorter, or longer than SEQ ID NO: 53, 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.

[00221] In some embodiments, Cas9 fusion proteins as provided herein comprise the full-length amino acid sequence of a Cas9 protein, *e.g.*, one of the Cas9 sequences provided herein. In other embodiments, however, fusion proteins as provided herein do not comprise a

full-length Cas9 sequence, but only a fragment thereof. For example, in some embodiments, a Cas9 fusion protein provided herein comprises a Cas9 fragment, wherein the fragment binds crRNA and tracrRNA or sgRNA, but does not comprise a functional nuclease domain, *e.g.*, in that it comprises only a truncated version of a nuclease domain or no nuclease domain at all.

[00222] Exemplary amino acid sequences of suitable Cas9 domains and Cas9 fragments are provided herein, and additional suitable sequences of Cas9 domains and fragments will be apparent to those of skill in the art.

[00223] In some embodiments, Cas9 refers to Cas9 from: *Corynebacterium ulcerans* (NCBI Refs: NC_015683.1, NC_017317.1); *Corynebacterium diphtheria* (NCBI Refs: NC_016782.1, NC_016786.1); *Spiroplasma syrphidicola* (NCBI Ref: NC_021284.1); *Prevotella intermedia* (NCBI Ref: NC_017861.1); *Spiroplasma taiwanense* (NCBI Ref: NC_021846.1); *Streptococcus iniae* (NCBI Ref: NC_021314.1); *Belliella baltica* (NCBI Ref: NC_018010.1); *Psychroflexus torquisI* (NCBI Ref: NC_018721.1); *Streptococcus thermophilus* (NCBI Ref: YP_820832.1); *Listeria innocua* (NCBI Ref: NP_472073.1); *Campylobacter jejuni* (NCBI Ref: YP_002344900.1); or *Neisseria. meningitidis* (NCBI Ref: YP_002342100.1).

[00224] It should be appreciated that additional Cas9 proteins (*e.g.,* a nuclease dead Cas9 (dCas9), a Cas9 nickase (nCas9), or a nuclease active Cas9), including variants and homologs thereof, are within the scope of this disclosure. Exemplary Cas9 proteins include, without limitation, those provided below. In some embodiments, the Cas9 protein is a nuclease dead Cas9 (dCas9). In some embodiments, the dCas9 comprises the amino acid sequence (SEQ ID NO: 34). In some embodiments, the Cas9 protein is a Cas9 nickase (nCas9). In some embodiments, the nCas9 comprises the amino acid sequence (SEQ ID NO: 35). In some embodiments, the Cas9 protein is a nuclease active Cas9. In some embodiments, the nuclease active Cas9 comprises the amino acid sequence (SEQ ID NO: 36).

Exemplary catalytically inactive Cas9 (dCas9):
DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL

FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:
34)

Exemplary Cas9 nickase (nCas9):

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN

FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:
35)

Exemplary catalytically active Cas9:
DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM

PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 36).

[00225] In some embodiments, Cas9 refers to a Cas9 from arehaea (*e.g.* nanoarchaea), which constitute a domain and kingdom of single-celled prokaryotic microbes. In some embodiments, Cas9 refers to CasX or 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–CasX and CRISPR–CasY, which are among the most compact systems yet discovered. In some embodiments, Cas9 refers to CasX, or a variant of CasX. In some embodiments, Cas9 refers to a CasY, or a variant of CasY. 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.

[00226] In some embodiments, the nucleic acid programmable DNA binding protein (napDNAbp) of any of the fusion proteins provided herein may be a CasX or CasY protein. In some embodiments, the napDNAbp is a CasX protein. In some embodiments, the napDNAbp is a 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 ease 99.5% identical to a naturally-occurring CasX or CasY protein. In some embodiments, the napDNAbp is a naturally-occurring CasX or 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 ease 99.5% identical to any one of SEQ ID NOs: 417-419. In some embodiments, the napDNAbp comprises an amino acid sequence of any one SEQ ID NOs:

417-419. It should be appreciated that CasX and CasY from other bacterial species may also be used in accordance with the present disclosure.

CasX (uniprot.org/uniprot/F0NN87; uniprot.org/uniprot/F0NH53)
>trlF0NN87lF0NN87_SULIH CRISPR-associated Casx protein OS=Sulfolobus islandicus (strain HVE10/4) GN=SiH_0402 PE=4 SV=1
MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIAKNNEDAAAE
RRGKAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFKNFSQV
KECEEVSAPSFVKPEFYEFGRSPGMVERTRRVKLEVEPHYLIIAAAGWVLTRLGKAK
VSEGDYVGVNVFTPTRGILYSLIQNVNGIVPGIKPETAFGLWIARKVVSSVTNPNVSV
VRIYTISDAVGQNPTTINGGFSIDLTKLLEKRYLLSERLEAIARNALSISSNMRERYIVL
ANYIYEYLTGSKRLEDLLYFANRDLIMNLNSDDGKVRDLKLISAYVNGELIRGEG
(SEQ ID NO: 417)

>trlF0NH53lF0NH53_SULIR CRISPR associated protein, Casx OS=Sulfolobus islandicus (strain REY15A) GN=SiRe_0771 PE=4 SV=1
MEVPLYNIFGDNYIIQVATEAENSTIYNNKVEIDDEELRNVLNLAYKIAKNNEDAAAE
RRGKAKKKKGEEGETTTSNIILPLSGNDKNPWTETLKCYNFPTTVALSEVFKNFSQV
KECEEVSAPSFVKPEFYKFGRSPGMVERTRRVKLEVEPHYLIMAAAGWVLTRLGKA
KVSEGDYVGVNVFTPTRGILYSLIQNVNGIVPGIKPETAFGLWIARKVVSSVTNPNVS
VVSIYTISDAVGQNPTTINGGFSIDLTKLLEKRDLLSERLEAIARNALSISSNMRERYIV
LANYIYEYLTGSKRLEDLLYFANRDLIMNLNSDDGKVRDLKLISAYVNGELIRGEG
(SEQ ID NO: 418)

CasY (ncbi.nlm.nih.gov/protein/APG80656.1)
>APG80656.1 CRISPR-associated protein CasY [uncultured Parcubacteria group bacterium]
MSKRHPRISGVKGYRLHAQRLEYTGKSGAMRTIKYPLYSSPSGGRTVPREIVSAINDD
YVGLYGLSNFDDLYNAEKRNEEKVYSVLDFWYDCVQYGAVFSYTAPGLLKNVAEV
RGGSYELTKTLKGSHLYDELQIDKVIKFLNKKEISRANGSLDKLKKDIIDCFKAEYRE
RHKDQCNKLADDIKNAKKDAGASLGERQKKLFRDFFGISEQSENDKPSFTNPLNLTC
CLLPFDTVNNNRNRGEVLFNKLKEYAQKLDKNEGSLEMWEYIGIGNSGTAFSNFLGE
GFLGRLRENKITELKKAMMDITDAWRGQEQEEELEKRLRILAALTIKLREPKFDNHW
GGYRSDINGKLSSWLQNYINQTVKIKEDLKGHKKDLKKAKEMINRFGESDTKEEAV
VSSLLESIEKIVPDDSADDEKPDIPAIAIYRRFLSDGRLTLNRFVQREDVQEALIKERLE

AEKKKKPKKRKKKSDAEDEKETIDFKELFPHLAKPLKLVPNFYGDSKRELYKKYKN
AAIYTDALWKAVEKIYKSAFSSSLKNSFFDTDFDKDFFIKRLQKIFSVYRRFNTDKWK
PIVKNSFAPYCDIVSLAENEVLYKPKQSRSRKSAAIDKNRVRLPSTENIAKAGIALARE
LSVAGFDWKDLLKKEEHEEYIDLIELHKTALALLLAVTETQLDISALDFVENGTVKD
FMKTRDGNLVLEGRFLEMFSQSIVFSELRGLAGLMSRKEFITRSAIQTMNGKQAELL
YIPHEFQSAKITTPKEMSRAFLDLAPAEFATSLEPESLSEKSLLKLKQMRYYPHYFGY
ELTRTGQGIDGGVAENALRLEKSPVKKREIKCKQYKTLGRGQNKIVLYVRSSYYQTQ
FLEWFLHRPKNVQTDVAVSGSFLIDEKKVKTRWNYDALTVALEPVSGSERVFVSQPF
TIFPEKSAEEEGQRYLGIDIGEYGIAYTALEITGDSAKILDQNFISDPQLKTLREEVKGL
KLDQRRGTFAMPSTKIARIRESLVHSLRNRIHHLALKHKAKIVYELEVSRFEEGKQKI
KKVYATLKKADVYSEIDADKNLQTTVWGKLAVASEISASYTSQFCGACKKLWRAE
MQVDETITTQELIGTVRVIKGGTLIDAIKDFMRPPIFDENDTPFPKYRDFCDKHHISKK
MRGNSCLFICPFCRANADADIQASQTIALLRYVKEEKKVEDYFERFRKLKNIKVLGQ
MKKI (SEQ ID NO: 419)

[00227] The term "effective amount," as used herein, refers to an amount of a biologically active agent that is sufficient to elicit a desired biological response. For example, in some embodiments, an effective amount of a nucleobase editor may refer to the amount of the nucleobase editor that is sufficient to induce mutation of a target site specifically bound mutated by the nucleobase editor. In some embodiments, an effective amount of a fusion protein provided herein, *e.g.*, of a fusion protein comprising a nucleic acid programmable DNA binding protein and a deaminase domain (*e.g.*, an adenosine deaminase domain) may refer to the amount of the fusion protein that is sufficient to induce editing of a target site specifically bound and edited by the fusion protein. As will be appreciated by the skilled artisan, the effective amount of an agent, *e.g.*, a fusion protein, a nucleobase editor, a deaminase, a hybrid protein, a protein dimer, a complex of a protein (or protein dimer) and a polynucleotide, or a polynucleotide, may vary depending on various factors as, for example, on the desired biological response, *e.g.*, on the specific allele, genome, or target site to be edited, on the cell or tissue being targeted, and on the agent being used.

[00228] The terms "nucleic acid" and "nucleic acid molecule," as used herein, refer to a compound comprising a nucleobase and an acidic moiety, *e.g.*, a nucleoside, a nucleotide, or a polymer of nucleotides. Typically, polymeric nucleic acids, *e.g.*, nucleic acid molecules comprising three or more nucleotides are linear molecules, in which adjacent nucleotides are linked to each other via a phosphodiester linkage. In some embodiments, "nucleic acid"

refers to individual nucleic acid residues (*e.g.* nucleotides and/or nucleosides). In some embodiments, "nucleic acid" refers to an oligonucleotide chain comprising three or more individual nucleotide residues. As used herein, the terms "oligonucleotide" and "polynucleotide" can be used interchangeably to refer to a polymer of nucleotides (*e.g.*, a string of at least three nucleotides). In some embodiments, "nucleic acid" encompasses RNA as well as single and/or double-stranded DNA. Nucleic acids may be naturally occurring, for example, in the context of a genome, a transcript, an mRNA, tRNA, rRNA, siRNA, snRNA, a plasmid, cosmid, chromosome, chromatid, or other naturally occurring nucleic acid molecule. On the other hand, a nucleic acid molecule may be a non-naturally occurring molecule, *e.g.*, a recombinant DNA or RNA, an artificial chromosome, an engineered genome, or fragment thereof, or a synthetic DNA, RNA, DNA/RNA hybrid, or including non-naturally occurring nucleotides or nucleosides. Furthermore, the terms "nucleic acid," "DNA," "RNA," and/or similar terms include nucleic acid analogs, *e.g.,* analogs having other than a phosphodiester backbone. Nucleic acids can be purified from natural sources, produced using recombinant expression systems and optionally purified, chemically synthesized, *etc.* Where appropriate, *e.g.*, in the case of chemically synthesized molecules, nucleic acids can comprise nucleoside analogs such as analogs having chemically modified bases or sugars, and backbone modifications. A nucleic acid sequence is presented in the 5′ to 3′ direction unless otherwise indicated. In some embodiments, a nucleic acid is or comprises natural nucleosides (*e.g.* 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, 2-aminoadenosine, C5-bromouridine, C5-fluorouridine, C5-iodouridine, C5-propynyl-uridine, C5-propynyl-cytidine, C5-methylcytidine, 2-aminoadenosine, 7-deazaadenosine, 7-deazaguanosine, 8-oxoadenosine, 8-oxoguanosine, O(6)-methylguanine, and 2-thiocytidine); chemically modified bases; biologically modified bases (*e.g.*, methylated bases); intercalated bases; modified sugars (*e.g.*, 2′-fluororibose, ribose, 2′-deoxyribose, arabinose, and hexose); and/or modified phosphate groups (*e.g.*, phosphorothioates and 5′-*N*-phosphoramidite linkages).

[00229] The term "proliferative disease," as used herein, refers to any disease in which cell or tissue homeostasis is disturbed in that a cell or cell population exhibits an abnormally elevated proliferation rate. Proliferative diseases include hyperproliferative diseases, such as pre-neoplastic hyperplastic conditions and neoplastic diseases. Neoplastic diseases are

characterized by an abnormal proliferation of cells and include both benign and malignant neoplasias. Malignant neoplasia is also referred to as cancer.

[00230] 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. 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. In some embodiments, a protein comprises a proteinaceous part, *e.g.*, an amino acid sequence constituting a nucleic acid binding domain, and an organic compound, *e.g.*, a compound that can act as a nucleic acid cleavage agent. In some embodiments, a protein is in a complex with, or is in association with, a nucleic acid, *e.g.*, RNA. 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.

[0001] The term "RNA-programmable nuclease," and "RNA-guided nuclease" are used interchangeably herein and refer to a nuclease that forms a complex with (*e.g.*, binds or

associates with) one or more RNA(s) that is not a target for cleavage. In some embodiments, an RNA-programmable nuclease, when in a complex with an RNA, may be referred to as a nuclease:RNA complex. Typically, the bound RNA(s) is referred to as a guide RNA (gRNA). gRNAs can exist as a complex of two or more RNAs, or as a single RNA molecule. gRNAs that exist as a single RNA molecule may be referred to as single-guide RNAs (sgRNAs), though "gRNA" is used interchangeably to refer to guide RNAs that exist as either single molecules or as a complex of two or more molecules. Typically, gRNAs that exist as single RNA species comprise two domains: (1) a domain that shares homology to a target nucleic acid (*e.g.*, and directs binding of a Cas9 complex to the target); and (2) a domain that binds a Cas9 protein. In some embodiments, domain (2) corresponds to a sequence known as a tracrRNA, and comprises a stem-loop structure. For example, in some embodiments, domain (2) is identical or homologous to a tracrRNA as provided in Jinek *et al.*, *Science* 337:816-821(2012), the entire contents of which is incorporated herein by reference. Other examples of gRNAs (*e.g.*, those including domain 2) can be found in U.S. Provisional Patent Application, U.S.S.N. 61/874,682, filed September 6, 2013, entitled "Switchable Cas9 Nucleases And Uses Thereof," and U.S. Provisional Patent Application, U.S.S.N. 61/874,746, filed September 6, 2013, entitled "Delivery System For Functional Nucleases," the entire contents of each are hereby incorporated by reference in their entirety. In some embodiments, a gRNA comprises two or more of domains (1) and (2), and may be referred to as an "extended gRNA." For example, an extended gRNA will, *e.g.*, bind two or more Cas9 proteins and bind a target nucleic acid at two or more distinct regions, as described herein. The gRNA comprises a nucleotide sequence that complements a target site, which mediates binding of the nuclease/RNA complex to said target site, providing the sequence specificity of the nuclease:RNA complex. In some embodiments, the RNA-programmable nuclease is the (CRISPR-associated system) Cas9 endonuclease, for example, Cas9 (Csn1) from *Streptococcus pyogenes* (see, *e.g.*, "Complete genome sequence of an M1 strain of *Streptococcus pyogenes*." Ferretti 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.

[00231] Because RNA-programmable nucleases (*e.g.*, Cas9) use RNA:DNA hybridization to target DNA cleavage sites, these proteins are able to be targeted, in principle, to any sequence specified by the guide RNA. Methods of using RNA-programmable nucleases, such as Cas9, for site-specific cleavage (*e.g.*, to modify a genome) are known in the art (see *e.g.*, Cong, L. *et al.*, Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819-823 (2013); Mali, P. *et al.*, RNA-guided human genome engineering via Cas9. *Science* **339**, 823-826 (2013); Hwang, W.Y. *et al.*, Efficient genome editing in zebrafish using a CRISPR-Cas system. *Nature biotechnology* **31**, 227-229 (2013); Jinek, M. *et al.*, RNA-programmed genome editing in human cells. *eLife* **2**, e00471 (2013); Dicarlo, J.E. *et al.*, Genome engineering in *Saccharomyces cerevisiae* using CRISPR-Cas systems. *Nucleic acids research* (2013); Jiang, W. *et al.* RNA-guided editing of bacterial genomes using CRISPR-Cas systems. *Nature biotechnology* **31**, 233-239 (2013); the entire contents of each of which are incorporated herein by reference).

[00232] The term "subject," as used herein, refers to an individual organism, for example, an individual mammal. In some embodiments, the subject is a human. In some embodiments, the subject is a non-human mammal. In some embodiments, the subject is a non-human primate. In some embodiments, the subject is a rodent. In some embodiments, the subject is a sheep, a goat, a cattle, a cat, or a dog. In some embodiments, the subject is a vertebrate, an amphibian, a reptile, a fish, an insect, a fly, or a nematode. In some embodiments, the subject is a research animal. In some embodiments, the subject is genetically engineered, *e.g.*, a genetically engineered non-human subject. The subject may be of either sex and at any stage of development.

[00233] The term "target site" refers to a sequence within a nucleic acid molecule that is deaminated by a deaminase or a fusion protein comprising a deaminase, (*e.g.*, a dCas9-adenosine deaminase fusion protein provided herein).

[00234] The terms "treatment," "treat," and "treating," refer to a clinical intervention aimed to reverse, alleviate, delay the onset of, or inhibit the progress of a disease or disorder, or one or more symptoms thereof, as described herein. As used herein, the terms "treatment," "treat," and "treating" refer to a clinical intervention aimed to reverse, alleviate, delay the onset of, or inhibit the progress of a disease or disorder, or one or more symptoms thereof, as described herein. In some embodiments, treatment may be administered after one or more symptoms have developed and/or after a disease has been diagnosed. In other embodiments,

treatment may be administered in the absence of symptoms, *e.g.*, to prevent or delay onset of a symptom or inhibit onset or progression of a disease. For example, treatment may be administered to a susceptible individual prior to the onset of symptoms (*e.g.*, in light of a history of symptoms and/or in light of genetic or other susceptibility factors). Treatment may also be continued after symptoms have resolved, for example, to prevent or delay their recurrence.

[00235] The term "recombinant" as used herein in the context of proteins or nucleic acids refers to proteins or nucleic acids that do not occur in nature, but are the product of human engineering. For example, in some embodiments, a recombinant protein or nucleic acid molecule comprises an amino acid or nucleotide sequence that comprises at least one, at least two, at least three, at least four, at least five, at least six, or at least seven mutations as compared to any naturally occurring sequence.

DETAILED DESCRIPTION OF THE INVENTION

[00236] Some aspects of this disclosure relate to proteins that deaminate the nucleobase adenine. This disclosure provides adenosine deaminase proteins that are capable of deaminating (*i.e.*, removing an amine group) adenine of a deoxyadenosine residue in deoxyribonucleic acid (DNA). For example, the adenosine deaminases provided herein are capable of deaminating adenine of a deoxyadenosine residue of DNA. It should be appreciated that there were no known adenosine deaminases capable of deaminating deoxyadenosine in DNA before the present invention. Other aspects of the disclosure provide fusion proteins that comprise an adenosine deaminase (*e.g.*, an adenosine deaminase that deaminates deoxyadenosine in DNA as described herein) and a domain (*e.g.*, a Cas9 or a Cpf1 protein) capable of binding to a specific nucleotide sequence. The deamination of an adenosine by an adenosine deaminase can lead to a point mutation, this process is referred to herein as nucleic acid editing. For example, the adenosine may be converted to an inosine residue, which typically base pairs with a cytosine residue. Such fusion proteins are useful *inter alia* for targeted editing of nucleic acid sequences. Such fusion proteins may be used for targeted editing of DNA *in vitro, e.g.*, for the generation of mutant cells or animals; for the introduction of targeted mutations, *e.g.*, for the correction of genetic defects in cells *ex vivo*, *e.g.*, in cells obtained from a subject that are subsequently re-introduced into the same or another subject; and for the introduction of targeted mutations *in vivo*, *e.g.*, the correction of genetic defects or the introduction of deactivating mutations in disease-associated genes in a subject. As an example, diseases that can be treated by making an A to G, or a T to C

mutation may be treated using the nucleobase editors provided herein. The invention provides deaminases, fusion proteins, nucleic acids, vectors, cells, compositions, methods, kits, systems, etc. that utilize the deaminases and nucleobase editors.

[00237] In some embodiments, the nucleobase editors provided herein can be made by fusing together one or more protein domains, thereby generating a fusion protein. In certain embodiments, the fusion proteins provided herein comprise one or more features that improve the base editing activity (*e.g.*, efficiency, selectivity, and specificity) of the fusion proteins. For example, the fusion proteins provided herein may comprise a Cas9 domain that has reduced nuclease activity. In some embodiments, the fusion proteins provided herein may have a Cas9 domain that does not have nuclease activity (dCas9), or a Cas9 domain that cuts one strand of a duplexed DNA molecule, referred to as a Cas9 nickase (nCas9). Without wishing to be bound by any particular theory, the presence of the catalytic residue (*e.g.*, H840) maintains the activity of the Cas9 to cleave the non-edited (*e.g.*, non-deaminated) strand containing a T opposite the targeted A. Mutation of the catalytic residue (*e.g.*, D10 to A10) of Cas9 prevents cleavage of the edited strand containing the targeted A residue. Such Cas9 variants are able to generate a single-strand DNA break (nick) at a specific location based on the gRNA-defined target sequence, leading to repair of the non-edited strand, ultimately resulting in a T to C change on the non-edited strand. In some embodiments, any of the fusion proteins provided herein further comprise an inhibitor of inosine base excision repair, for example, a uracil glycosylase inhibitor (UGI) domain or a catalytically inactive inosine-specific nuclease (dISN). Without wishing to be bound by any particular theory, the UGI domain or dISN may inhibit or prevent base excision repair of a deaminated adenosine residue (*e.g.*, inosine), which may improve the activity or efficiency of the base editor.

Adenosine deaminases

[00238] Some aspects of the disclosure provide adenosine deaminases. In some embodiments, the adenosine deaminases provided herein are capable of deaminating adenine. In some embodiments, the adenosine deaminases provided herein are capable of deaminating adenine in a deoxyadenosine residue of DNA. The adenosine deaminase may be derived from any suitable organism (*e.g.*, *E. coli*). In some embodiments, the adenine deaminase is a naturally-occurring adenosine deaminase that includes one or more mutations corresponding to any of the mutations provided herein (*e.g.*, mutations in ecTadA). One of skill in the art will be able to identify the corresponding residue in any homologous protein and in the respective encoding nucleic acid by methods well known in the art, *e.g.*, by sequence

alignment and determination of homologous residues. Accordingly, one of skill in the art would be able to generate mutations in any naturally-occurring adenosine deaminase (*e.g.*, having homology to ecTadA) that corresponds to any of the mutations described herein, *e.g.*, any of the mutations identified in ecTadA. In some embodiments, the adenosine deaminase is from a prokaryote. In some embodiments, the adenosine deaminase is from a bacterium. In some embodiments, the adenosine deaminase is from *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi, Shewanella putrefaciens, Haemophilus influenzae, Caulobacter crescentus,* or *Bacillus subtilis.* In some embodiments, the adenosine deaminase is from *E. coli*.

[00239] Exemplary alignment of prokaryotic TadA proteins is shown in Figure 92. The residues highlighted in blue are the residues which may be important for catalyzing A to I deamination on ssDNA. Accordingly, it should be appreciated that any of the mutations identified in ecTadA provided herein may be made in any homologous residue in another adenine deaminase, for example, a TadA deaminase from another bacterium. Figure 93 shows the relative sequence identity analysis (heatmap of sequence identity) :

[00240] In some embodiments, the adenosine deaminase comprises an amino acid sequence that is 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% identical to any one of the amino acid sequences set forth in any one of SEQ ID NOs: 1, 64-84, 420-437, 672-684, or to any of the adenosine deaminases provided herein. It should be appreciated that adenosine deaminases provided herein may include one or more mutations (*e.g.*, any of the mutations provided herein). The disclosure provides any deaminase domains with a certain percent identiy plus any of the mutations or combinations thereof described herein. In some embodiments, the adenosine deaminase comprises an amino acid sequence that has 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 mutations compared to any one of the amino acid sequences set forth in SEQ ID NOs: 1, 64-84, 420-437, 672-684, or any of the adenosine deaminases provided herein. In some embodiments, the adenosine deaminase comprises an amino acid sequence that has at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 110, at least 120, at least 130, at least 140, at least 150, at least 160, or at least 170 identical contiguous amino acid residues as compared to any one of the amino acid

sequences set forth in SEQ ID NOs: 1, 64-84, 420-437, 672-684, or any of the adenosine deaminases provided herein.

Evolution #1 and #2 Mutations

[00241] In some embodiments, the adenosine deaminase comprises a D108X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a D108G, D108N, D108V, D108A, or D108Y mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase. An exemplary alignment of deaminases is shown in Figure 92. It should be appreciated, however, that additional deaminases may similarly be aligned to identify homologous amino acid residues that can be mutated as provided herein.

[00242] In some embodiments, the adenosine deaminse comprises an A106X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an A106V mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00243] In some embodiments, the adenosine deaminase comprises a E155X mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a E155D, E155G, or E155V mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00244] In some embodiments, the adenosine deaminase comprises a D147X mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a D147Y, mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00245] It should be appreciated that any of the mutations provided herein (*e.g.,* based on the ecTadA amino acid sequence of SEQ ID NO: 1) may be introduced into other adenosine deaminases, such as *S. aureus* TadA (saTadA), or other adenosine deaminases (*e.g.,* bacterial adenosine deaminases). It would be apparent to the skilled artisan how to identify amino acid residues from other adenosine deaminases that are homologous to the

mutated residues in ecTadA. Thus, any of the mutations identified in ecTadA may be made in other adenosine deaminases that have homologous amino acid residues. It should also be appreciated that any of the mutations provided herein may be made individually or in any combination in ecTadA or another adenosine deaminase. For example, an adenosine deaminase may contain a D108N, a A106V, a E155V, and/or a D147Y mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase. In some embodiments, an adenosine deaminase comprises the following group of mutations (groups of mutations are separated by a ";") in ecTadA SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase:
D108N and A106V; D108N and E155V; D108N and D147Y; A106V and E155V; A106V and D147Y; E155V and D147Y; D108N, A106V, and E55V; D108N, A106V, and D147Y; D108N, E55V, and D147Y; A106V, E55V, and D147Y; and D108N, A106V, E55V, and D147Y. It should be appreciated, however, that any combination of corresponding mutations provided herein may be made in an adenosine deaminase (*e.g.,* ecTadA). In some embodiments, an adenosine deaminase comprises one or more of the mutations shown in Table 4, which identifies individual mutations and combinations of mutations made in ecTadA and saTadA. In some embodiments, an adenosine deaminase comprises a mutation or combination of mutations shown in Table 4.

[00246] In some embodiments, the adenosine deaminase comprises one or more of a H8X, T17X, L18X, W23X, L34X, W45X, R51X, A56X, E59X, E85X, M94X, I95X, V102X, F104X, A106X, R107X, D108X, K110X, M118X, N127X, A138X, F149X, M151X, R153X, Q154X, I156X, and/or K157X mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of H8Y, T17S, L18E, W23L, L34S, W45L, R51H, A56E, or A56S, E59G, E85K, or E85G, M94L, I95I, V102A, F104L, A106V, R107C, or R107H, or R107P, D108G, or D108N, or D108V, or D108A, or D108Y, K110I, M118K, N127S, A138V, F149Y, M151V, R153C, Q154L, I156D, and/or K157R mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of the mutations provided in Figure 11 corresponding to SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises the mutation or mutations of any one of constructs 1-16

shown in Figure 11 or in any one of the constructs shown in Table 4 corresponding to SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase.

[00247] In some embodiments, the adenosine deaminase comprises one or more of a H8X, D108X, and/or N127X mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase, where X indicates the presence of any amino acid. In some embodiments, the adenosine deaminase comprises one or more of a H8Y, D108N, and/or N127S mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase.

[00248] In some embodiments, the adenosine deaminase comprises one or more of H8X, R26X, M61X, L68X, M70X, A106X, D108X, A109X, N127X, D147X, R152X, Q154X, E155X, K161X, Q163X, and/or T166X mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of H8Y, R26W, M61I, L68Q, M70V, A106T, D108N, A109T, N127S, D147Y, R152C, Q154H or Q154R, E155G or E155V or E155D, K161Q, Q163H, and/or T166P mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase.

[00249] In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of H8X, D108X, N127X, D147X, R152X, and Q154X in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, seven, or eight mutations selected from the group consisting of H8X, M61X, M70X, D108X, N127X, Q154X, E155X, and Q163X in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five, mutations selected from the group consisting of H8X, D108X, N127X, E155X, and T166X in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of H8X, A106X, D108X, N127X, E155X, and K161X in SEQ ID NO: 1, or a corresponding mutation or mutations in

another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, seven, or eight mutations selected from the group consisting of H8X, R126X, L68X, D108X, N127X, D147X, and E155X in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five, mutations selected from the group consisting of H8X, D108X, A109X, N127X, and E155X in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase.

[00250] In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of H8Y, D108N, N127S, D147Y, R152C, and Q154H in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, seven, or eight mutations selected from the group consisting of H8Y, M61I, M70V, D108N, N127S, Q154R, E155G and Q163H in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five, mutations selected from the group consisting of H8Y, D108N, N127S, E155V, and T166P in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of H8Y, A106T, D108N, N127S, E155D, and K161Q in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, seven, or eight mutations selected from the group consisting of H8Y, R126W, L68Q, D108N, N127S, D147Y, and E155V in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five, mutations selected from the group consisting of H8Y, D108N, A109T, N127S, and E155G in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase.

[00251] In some embodiments, the adenosine deaminase comprises one or more of the mutations provided in Figure 16 corresponding to SEQ ID NO: 1, or one or more

corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises the mutations of any one of constructs pNMG-149 to pNMG-154 of Figure 16, corresponding to SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a D108N, D108G, or D108V mutation in SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a A106V and D108N mutation in SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises R107C and D108N mutations in SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a H8Y, D108N, N127S, D147Y, and Q154H mutation in SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a H8Y, R24W, D108N, N127S, D147Y, and E155V mutation in SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a D108N, D147Y, and E155V mutation in SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a H8Y, D108N, and S127S mutation in SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises a A106V, D108N, D147Y and E155V mutation in SEQ ID NO: 1, or corresponding mutations in another adenosine deaminase.

[00252] In some embodiments, the adenosine deaminase comprises one or more of a, S2X, H8X, I49X, L84X, H123X, N127X, I156X and/or K160X mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of S2A, H8Y, I49F, L84F, H123Y, N127S, I156F and/or K160S mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of the mutations provided in Figure 97 corresponding to SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises the mutation or mutations of any one of clones 1-3 shown in Figure 97 corresponding to SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase.

[00253] In some embodiments, the adenosine deaminse comprises an L84X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase,

where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an L84F mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00254] In some embodiments, the adenosine deaminse comprises an H123X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an H123Y mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00255] In some embodiments, the adenosine deaminse comprises an I157X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an I157F mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00256] In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, or seven mutations selected from the group consisting of L84X, A106X, D108X, H123X, D147X, E155X, and I156X in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of S2X, I49X, A106X, D108X, D147X, and E155X in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, or five, mutations selected from the group consisting of H8X, A106X, D108X, N127X, and K160X in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase, where X indicates the presence of any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase.

[00257] In some embodiments, the adenosine deaminase comprises one, two, three, four, five, six, or seven mutations selected from the group consisting of L84F, A106V, D108N, H123Y, D147Y, E155V, and I156F in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one, two, three, four, five, or six mutations selected from the group consisting of S2A, I49F, A106V, D108N, D147Y, and E155V in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase. In some

embodiments, the adenosine deaminase comprises one, two, three, four, or five, mutations selected from the group consisting of H8Y, A106T, D108N, N127S, and K160S in SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase.

[00258] In some embodiments, the adenosine deaminase comprises one or more of a, E25X, R26X, R107X, A142X, and/or A143X mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of E25M, E25D, E25A, E25R, E25V, E25S, E25Y, R26G, R26N, R26Q, R26C, R26L, R26K, R107P, R07K, R107A, R107N, R107W, R107H, R107S, A142N, A142D, A142G, A143D, A143G, A143E, A143L, A143W, A143M, A143S, A143Q and/or A143R mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of the mutations provided in Table 7 corresponding to SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises the mutation or mutations of any one of clones 1-22 shown in Table 7 corresponding to SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase.

[00259] In some embodiments, the adenosine deaminse comprises an E25X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an E25M, E25D, E25A, E25R, E25V, E25S, or E25Y mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00260] In some embodiments, the adenosine deaminse comprises an R26X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an , R26G, R26N, R26Q, R26C, R26L, or R26K mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00261] In some embodiments, the adenosine deaminse comprises an R107X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an R107P,

R07K, R107A, R107N, R107W, R107H, or R107S mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00262] In some embodiments, the adenosine deaminse comprises an A142X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an A142N, A142D, A142G, mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00263] In some embodiments, the adenosine deaminse comprises an A143X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an A143D, A143G, A143E, A143L, A143W, A143M, A143S, A143Q and/or A143R mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00264] In some embodiments, the adenosine deaminase comprises one or more of a, H36X, N37X, P48X, I49X, R51X, M70X, N72X, D77X, E134X, S146X, Q154X, K157X, and/or K161X mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase, where the presence of X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of H36L, N37T, N37S, P48T, P48L, I49V, R51H, R51L, M70L, N72S, D77G, E134G, S146R, S146C, Q154H, K157N, and/or K161T mutation in SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises one or more of the mutations provided in any one of Figures 125-128 corresponding to SEQ ID NO: 1, or one or more corresponding mutations in another adenosine deaminase. In some embodiments, the adenosine deaminase comprises the mutation or mutations of any one of clones 1-11 shown in any one of Figures 125-128 corresponding to SEQ ID NO: 1, or a corresponding mutation or mutations in another adenosine deaminase.

[00265] In some embodiments, the adenosine deaminse comprises an H36X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an H36L mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00266] In some embodiments, the adenosine deaminse comprises an N37X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an N37T, or N37S mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00267] In some embodiments, the adenosine deaminse comprises an P48X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an P48T, or P48L mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00268] In some embodiments, the adenosine deaminse comprises an R51X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an R51H, or R51L mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00269] In some embodiments, the adenosine deaminse comprises an S146X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises an S146R, or S146C mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00270] In some embodiments, the adenosine deaminse comprises an K157X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a K157N mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00271] In some embodiments, the adenosine deaminse comprises an P48X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a P48S,

P48T, or P48A mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00272] In some embodiments, the adenosine deaminse comprises an A142X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a A142N mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00273] In some embodiments, the adenosine deaminse comprises an W23X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a W23R, or W23L mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00274] In some embodiments, the adenosine deaminse comprises an R152X mutation in ecTadA SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase, where X indicates any amino acid other than the corresponding amino acid in the wild-type adenosine deaminase. In some embodiments, the adenosine deaminase comprises a R152P, or R52H mutation in SEQ ID NO: 1, or a corresponding mutation in another adenosine deaminase.

[00275] It should be appreciated that the adenosine deaminase (*e.g.,* a first or second adenosine deaminase) may comprise one or more of the mutations provided in any of the adenosine deaminases (*e.g.,* ecTadA adenosine deaminases) shown in Table 4. In some embodiments, the adenosine deaminase comprises the combination of mutations of any of the adenosine deaminases (*e.g.,* ecTadA adenosine deaminases) shown in Table 4. For example, the adenosine deaminase may comprise the mutations H36L, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V, I156F, and K157N, which are shown in the second ecTadA (relative to SEQ ID NO: 1) of clone pNMG-477. In some embodiments, the adenosine deaminase comprises the following combination of mutations relative to SEQ ID NO:1, where each mutation of a combination is separated by a "_" and each combination of mutations is between parentheses: (A106V_D108N), (R107C_D108N), (H8Y_D108N_S127S_D147Y_Q154H), (H8Y_R24W_D108N_N127S_D147Y_E155V), (D108N_D147Y_E155V), (H8Y_D108N_S127S), (H8Y_D108N_N127S_D147Y_Q154H), (A106V_D108N_D147Y_E155V), (D108Q_D147Y_E155V), (D108M_D147Y_E155V),

(D108L_D147Y_E155V), (D108K_D147Y_E155V), (D108I_D147Y_E155V), (D108F_D147Y_E155V), (A106V_D108N_D147Y), (A106V_D108M_D147Y_E155V), (E59A_A106V_D108N_D147Y_E155V), (E59A cat dead_A106V_D108N_D147Y_E155V), (L84F_A106V_D108N_H123Y_D147Y_E155V_I156Y), (L84F_A106V_D108N_H123Y_D147Y_E155V_I156F), (D103A_D014N), (G22P_D103A_D104N), (G22P_D103A_D104N_S138A), (D103A_D104N_S138A), (R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (E25G_R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (E25D_R26G_L84F_A106V_R107K_D108N_H123Y_A142N_A143G_D147Y_E155V_I156F), (R26Q_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (E25M_R26G_L84F_A106V_R107P_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (R26C_L84F_A106V_R107H_D108N_H123Y_A142N_D147Y_E155V_I156F), (L84F_A106V_D108N_H123Y_A142N_A143L_D147Y_E155V_I156F), (R26G_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (E25A_R26G_L84F_A106V_R107N_D108N_H123Y_A142N_A143E_D147Y_E155V_I156F), (R26G_L84F_A106V_R107H_D108N_H123Y_A142N_A143D_D147Y_E155V_I156F), (A106V_D108N_A142N_D147Y_E155V), (R26G_A106V_D108N_A142N_D147Y_E155V), (E25D_R26G_A106V_R107K_D108N_A142N_A143G_D147Y_E155V), (R26G_A106V_D108N_R107H_A142N_A143D_D147Y_E155V), (E25D_R26G_A106V_D108N_A142N_D147Y_E155V), (A106V_R107K_D108N_A142N_D147Y_E155V), (A106V_D108N_A142N_A143G_D147Y_E155V), (A106V_D108N_A142N_A143L_D147Y_E155V), (H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N), (N37T_P48T_M70L_L84F_A106V_D108N_H123Y_D147Y_I49V_E155V_I156F), (N37S_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K161T), (H36L_L84F_A106V_D108N_H123Y_D147Y_Q154H_E155V_I156F), (N72S_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F), (H36L_P48L_L84F_A106V_D108N_H123Y_E134G_D147Y_E155V_I156F), (H36L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N),

(H36L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F_K161T),
(N37S_R51H_D77G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(R51L_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_K157N),
(D24G_Q71R_L84F_H96L_A106V_D108N_H123Y_D147Y_E155V_I156F_K160E),
(H36L_G67V_L84F_A106V_D108N_H123Y_S146T_D147Y_E155V_I156F),
(Q71L_L84F_A106V_D108N_H123Y_L137M_A143E_D147Y_E155V_I156F),
(E25G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L),
(L84F_A91T_F104I_A106V_D108N_H123Y_D147Y_E155V_I156F),
(N72D_L84F_A106V_D108N_H123Y_G125A_D147Y_E155V_I156F),
(P48S_L84F_S97C_A106V_D108N_H123Y_D147Y_E155V_I156F),
(W23G_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(D24G_P48L_Q71R_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F_Q159L),
(L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F),
(H36L_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F
_K157N), (N37S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_K161T),
(L84F_A106V_D108N_D147Y_E155V_I156F),
(R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E_K161T),
(L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N_K160E), (R74Q
L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(R74A_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(R74Q_L84F_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_R98Q_A106V_D108N_H123Y_D147Y_E155V_I156F),
(L84F_A106V_D108N_H123Y_R129Q_D147Y_E155V_I156F),
(P48S_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F), (P48S_A142N),
(P48T_I49V_L84F_A106V_D108N_H123Y_A142N_D147Y_E155V_I156F_L157N),
(P48T_I49V_A142N),
(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F
_K157N),
(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F
_K157N),

(H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),

(H36L_P48T_I49V_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_ I156F _K157N),

(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),

(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_E155V_I156F _K157N),

(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F _K157N),

(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),

(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F _K157N),

(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F _K161T),

(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152H_E155V_I156F _K157N),

(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N),

(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V _I156F _K157N),

(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_E155 V_I156F _K157N),

(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_R152P _E155V_I156F _K157N),

(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146R_D147Y_E155V_I156F _K161T),

(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V _I156F _K157N),

(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152P_E155 V_I156F _K157N).

[00276] In some embodiments, the adenosine deaminase comprises an amino acid sequence that is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95, 98%, 99%, or 99.5%

identical to any one of SEQ ID NOs: 1, 64-84, 420-437, 672-684, or any of the adenosine deaminases provided herein. In some embodiments, the adenosine deaminase comprises an amino acid sequence that has 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 mutations compared to any one of the amino acid sequences set forth in SEQ ID NOs: 1, 64-84, 420-437, 672-684, or any of the adenosine deaminases provided herein. In some embodiments, the adenosine deaminase comprises an amino acid sequence that has at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 110, at least 120, at least 130, at least 140, at least 150, at least 160, or at least 166, identical contiguous amino acid residues as compared to any one of the amino acid sequences set forth in SEQ ID NOs: 1, 64-84, 420-437, 672-684, or any of the adenosine deaminases provided herein. In some embodiments, the adenosine deaminase comprises the amino acid sequence of any one of SEQ ID NOs: 1, 64-84, 420-437, 672-684, or any of the adenosine deaminases provided herein. In some embodiments, the adenosine deaminase consists of the amino acid sequence of any one of SEQ ID NOs: 1, 64-84, 420-437, 672-684, or any of the adenosine deaminases provided herein. The ecTadA sequences provided below are from ecTadA (SEQ ID NO: 1), absent the N-terminal methionine (M). The saTadA sequences provided below are from saTadA (SEQ DI NO: 8), absent the N-terminal methionine (M). For clarity, the amino acid numbering scheme used to identify the various amino acid mutations is derived from ecTadA (SEQ ID NO: 1) for *E. coli* TadA and saTadA (SEQ ID NO: 8) for *S. aureus* TadA. Amino acid mutations, relative to SEQ ID NO: 1 (ecTadA) or SEQ DI NO: 8 (saTadA), are indicated by underlining.

ecTadA
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 64)

ecTadA (D108N)
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGAR<u>N</u>AKTG

AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 65)

ecTadA (D108G)
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARGAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 66)

ecTadA (D108V)
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARVAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 67)

ecTadA (H8Y, D108N, and N127S)
SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 68)

ecTadA (H8Y, D108N, N127S, and E155D)
SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQDIKAQKKAQSSTD
(SEQ ID NO: 69)

ecTadA (H8Y, D108N, N127S, and E155G)
SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQGIKAQKKAQSSTD
(SEQ ID NO: 70)

ecTadA (H8Y, D108N, N127S, and E155V)

SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARNAKTG
AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQVIKAQKKAQSSTD
(SEQ ID NO: 71)

ecTadA (A106V, D108N, D147Y, and E155V)
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTD
(SEQ ID NO: 72)

ecTadA (L84F, A106V, D108N, H123Y, D147Y, E155V, I156F) - result of evolution #3
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 73)

ecTadA (S2A, I49F, A106V, D108N, D147Y, E155V) - result of evolution #3
AEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPFGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTD
(SEQ ID NO: 74)

ecTadA (H8Y, A106T, D108N, N127S, K160S) - result of evolution #3
SEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGTRNAKTG
AAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKAQSKAQSSTD
(SEQ ID NO: 75)

ecTadA (R26G, L84F, A106V, R107H, D108N, H123Y, A142N, A143D, D147Y, E155V, I156F) - result of evolution #4
SEVEFSHEYWMRHALTLAKRAWDEGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTG

AAGSLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 76)

ecTadA (E25G, R26G, L84F, A106V, R107H, D108N, H123Y, A142N, A143D, D147Y, E155V, I156F) - result of evolution #4
SEVEFSHEYWMRHALTLAKRAWDGGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 77)

ecTadA (E25D, R26G, L84F, A106V, R107K, D108N, H123Y, A142N, A143G, D147Y, E155V, I156F) -result of evolution #4
SEVEFSHEYWMRHALTLAKRAWDDGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVKNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNGLLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 78)

ecTadA (R26Q, L84F, A106V, D108N, H123Y, A142N, D147Y, E155V, I156F) - result of evolution #4
SEVEFSHEYWMRHALTLAKRAWDEQEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 79)

ecTadA (E25M, R26G, L84F, A106V, R107P, D108N, H123Y, A142N, A143D, D147Y, E155V, I156F) - result of evolution #4
SEVEFSHEYWMRHALTLAKRAWDMGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVPNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNDLLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 80)

ecTadA (R26C, L84F, A106V, R107H, D108N, H123Y, A142N , D147Y, E155V, I156F) - result of evolution #4

SEVEFSHEYWMRHALTLAKRAWDECEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVHNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 81)

ecTadA (L84F, A106V , D108N, H123Y, A142N, A143L, D147Y, E155V, I156F) - result of evolution #4
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNLLLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 82)

ecTadA (R26G, L84F, A106V, D108N, H123Y, A142N , D147Y, E155V, I156F) - result of evolution #4
SEVEFSHEYWMRHALTLAKRAWDEGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 83)

ecTadA (E25A, R26G, L84F, A106V, R107N, D108N, H123Y, A142N, A143E, D147Y, E155V, I156F) - result of evolution #4
SEVEFSHEYWMRHALTLAKRAWDAGEVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVNNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECNELLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 420)

ecTadA (L84F, A106V, D108N, H123Y, D147Y, E155V, I156F) – mutations from evolution #'s 1-3
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 421)

ecTadA (N37T, P48T, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F) – mutations from evolution # 5-1
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHTNRVIGEGWNRTIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 422)

ecTadA (N37S, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F) – mutations from evolution # 5-2
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHSNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 423)

ecTadA (H36L, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F) – mutations from evolution # 5-3
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 424)

ecTadA (L84F, A106V, D108N, H123Y, S146R, D147Y, E155V, I156F) – mutations from evolution # 5-4
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLRYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 425)

ecTadA (H36L, P48L, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F) – mutations from evolution # 5-5
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRLIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 426)

ecTadA (H36L, L84F, A106V, D108N, H123Y, D147Y, E155V, K57N, I156F) – mutations from evolution # 5-6
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 427)

ecTadA (H36L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V, I156F) – mutations from evolution # 5-7
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 428)

ecTadA (L84F, A106V, D108N, H123Y, S146R, D147Y, E155V, I156F) – mutations from evolution # 5-8
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLRYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 429)

ecTadA (N37S, R51H, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F) – mutations from evolution # 5-9
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHSNRVIGEGWNRPIGHHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 430)

ecTadA (R51L, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F, K157N) – mutations from evolution # 5-10
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD

(SEQ ID NO: 431)

ecTadA (R51H, L84F, A106V, D108N, H123Y, D147Y, E155V, I156F, K157N) – mutations from evolution # 5-11
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGHHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 432)

saTadA (wt) - as used in pNMG-345:
MGSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAH
AEHIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADDPKGGCS
GSLMNLLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 8)

saTadA (D108N) – as used in pNMG-346:
GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGADNPKGGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 433)

saTadA (D107A_D108N) – as used in pNMG-347:
GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 434)

saTadA (G26P_D107A_D108N) – as used in pNMG-348:
GSHMTNDIYFMTLAIEEAKKAAQLPEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACSTLLTTFFKNLRANKKSTN (SEQ ID NO: 435)

saTadA (G26P_D107A_D108N_S142A) – as used in pNMG-349:
GSHMTNDIYFMTLAIEEAKKAAQLPEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACATLLTTFFKNLRANKKSTN (SEQ ID NO: 436)

saTadA (D107A_D108N_S142A) – as used in pNMG-350:
GSHMTNDIYFMTLAIEEAKKAAQLGEVPIGAIITKDDEVIARAHNLRETLQQPTAHAE
HIAIERAAKVLGSWRLEGCTLYVTLEPCVMCAGTIVMSRIPRVVYGAANPKGGCSGS
LMNLLQQSNFNHRAIVDKGVLKEACATLLTTFFKNLRANKKSTN (SEQ ID NO: 437)

ecTadA (P48S) – mutation from evolution #6
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRSIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 672)

ecTadA (P48T) – mutation from evolution #6
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRTIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 673)

ecTadA (P48A) – mutation from evolution #6
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRAIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 674)

ecTadA (A142N) – mutation from evolution #6
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECNALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 675)

ecTadA (W23R) – mutation from evolution #7
SEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 676)

ecTadA (W23L) – mutation from evolution #7
SEVEFSHEYWMRHALTLAKRALDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTGA
AGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTD
(SEQ ID NO: 677)

ecTadA (R152P) – mutation from evolution #7
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMPRQEIKAQKKAQSSTD
(SEQ ID NO: 678)

ecTadA (R152H) – mutation from evolution #7
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGARDAKTG
AAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMHRQEIKAQKKAQSSTD
(SEQ ID NO: 679)

ecTadA (L84F, A106V, D108N, H123Y, D147Y, E155V, I156F) – mutations from pNMG 371
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD
(SEQ ID NO: 680)

ecTadA (H36L, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V, I156F, K157N) – mutations from pNMG 477
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 681)

ecTadA (H36L, P48S, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V, I156F , K157N) – mutations from pNMG 576
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRSIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 682)

ecTadA (H36L, P48A, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y, E155V, I156F , K157N) – mutations from pNMG 586
SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTA
HAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTG
AAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD
(SEQ ID NO: 683)

ecTadA (W23L, H36L, P48A, R51L, L84F, A106V, D108N, H123Y, S146C, D147Y, R152P, E155V, I156F, K157N) – mutations from pNMG 616
SEVEFSHEYWMRHALTLAKRALDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAH
AEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGA
AGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD
(SEQ ID NO: 684)

Cas9 Domains of Nucleobase Editors

[00277] In some aspects, a nucleic acid programmable DNA binding protein (napDNAbp) is a Cas9 domain. Non-limiting, exemplary Cas9 domains are provided herein. The Cas9 domain may be a nuclease active Cas9 domain, a nuclease inactive Cas9 domain, or a Cas9 nickase. In some embodiments, the Cas9 domain is a nuclease active domain. For example, the Cas9 domain may be a Cas9 domain that cuts both strands of a duplexed nucleic acid (*e.g.,* both strands of a duplexed DNA molecule). In some embodiments, the Cas9 domain comprises any one of the amino acid sequences as set forth in SEQ ID NOs: 108-357. In some embodiments the Cas9 domain comprises an amino acid sequence that is 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% identical to any one of the amino acid sequences set forth in SEQ ID NOs: 108-357. In some embodiments, the Cas9 domain comprises an amino acid sequence that has 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 or more mutations compared to any one of the amino acid sequences set forth in SEQ ID NOs: 108-357. In some embodiments, the Cas9 domain comprises an amino acid sequence that has at least 10, at least 15, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250, at least 300, at least 350, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, or at least 1200 identical contiguous amino acid residues as compared to any one of the amino acid sequences set forth in SEQ ID NOs: 108-357.

[00278] In some embodiments, the Cas9 domain is a nuclease-inactive Cas9 domain (dCas9). For example, the dCas9 domain may bind to a duplexed nucleic acid molecule (*e.g.*, via a gRNA molecule) without cleaving either strand of the duplexed nucleic acid molecule. In some embodiments, the nuclease-inactive dCas9 domain comprises a D10X mutation and a H840X mutation of the amino acid sequence set forth in SEQ ID NO: 52, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357, wherein X is any amino acid change. In some embodiments, the nuclease-inactive dCas9 domain comprises a D10A mutation and a H840A mutation of the amino acid sequence set forth in SEQ ID NO: 52, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357. As one example, a nuclease-inactive Cas9 domain comprises the amino acid sequence set forth in SEQ ID NO: 54 (Cloning vector pPlatTET-gRNA2, Accession No. BAV54124).

MDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGE
TAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHE
RHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEG
DLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLP
GEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYA
DLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPE
KYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQR
TFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFA
WMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTV
YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFD
SVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN

FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDAIVPQSFLKDDSIDNKVLTRSDK NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 54; see, *e.g.*, Qi *et al.*, "Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression." *Cell.* 2013; 152(5):1173-83, the entire contents of which are incorporated herein by reference).

[00279] Additional suitable nuclease-inactive dCas9 domains will be apparent to those of skill in the art based on this disclosure and knowledge in the field, and are within the scope of this disclosure. Such additional exemplary suitable nuclease-inactive Cas9 domains include, but are not limited to, D10A/H840A, D10A/D839A/H840A, and D10A/D839A/H840A/N863A mutant domains (See, *e.g.*, Prashant *et al.*, CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. *Nature Biotechnology*. 2013; 31(9): 833-838, the entire contents of which are incorporated herein by reference). In some embodiments the dCas9 domain comprises an amino acid sequence that is 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% identical to any one of the dCas9 domains provided herein. In some embodiments, the Cas9 domain comprises an amino acid sequences that has 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 mutations compared to any one of the amino acid sequences set forth in SEQ ID NOs: 108-357. In some embodiments, the Cas9 domain comprises an amino acid sequence that has at least 10, at least 15, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 150, at least 200, at least 250, at least 300, at least 350, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least

1000, at least 1100, or at least 1200 identical contiguous amino acid residues as compared to any one of the amino acid sequences set forth in SEQ ID NOs: 108-357.

[00280] In some embodiments, the Cas9 domain is a Cas9 nickase. The Cas9 nickase may be a Cas9 protein that is capable of cleaving only one strand of a duplexed nucleic acid molecule (*e.g.,* a duplexed DNA molecule). In some embodiments the Cas9 nickase cleaves the target strand of a duplexed nucleic acid molecule, meaning that the Cas9 nickase cleaves the strand that is base paired to (complementary to) a gRNA (*e.g.,* an sgRNA) that is bound to the Cas9. In some embodiments, a Cas9 nickase comprises a D10A mutation and has a histidine at position 840 of SEQ ID NO: 52, or a mutation in any of SEQ ID NOs: 108-357. As one example, a Cas9 nickase may comprise the amino acid sequence as set forth in SEQ ID NO: 35. In some embodiments, the Cas9 nickase cleaves the non-target, non-base-edited strand of a duplexed nucleic acid molecule, meaning that the Cas9 nickase cleaves the strand that is not base paired to a gRNA (*e.g.,* an sgRNA) that is bound to the Cas9. In some embodiments, a Cas9 nickase comprises an H840A mutation and has an aspartic acid residue at position 10 of SEQ ID NO: 52, or a corresponding mutation in any of SEQ ID NOs: 108-357. In some embodiments the Cas9 nickase comprises an amino acid sequence that is 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% identical to any one of the Cas9 nickases provided herein. Additional suitable Cas9 nickases will be apparent to those of skill in the art based on this disclosure and knowledge in the field, and are within the scope of this disclosure.

Cas9 Domains with Reduced PAM Exclusivity

[00281] 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, where the "N" in "NGG" is adenine (A), thymine (T), guanine (G), or cytosine (C), and the G is guanine. This may limit the ability to edit desired bases within a genome. In some embodiments, the base editing fusion proteins provided herein need to be positioned at a precise location, for example, where a target base is within a 4 base region (*e.g.,* a "deamination 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. In some embodiments, the deamination window is within a 2, 3, 4, 5, 6, 7, 8, 9, or 10 base

region. In some embodiments, the deamination window is 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, or 25 bases upstream of the PAM. 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.

[00282] In some embodiments, the Cas9 domain is a Cas9 domain from *Staphylococcus aureus* (SaCas9). In some embodiments, the SaCas9 domain is a nuclease active SaCas9, a nuclease inactive SaCas9 (SaCas9d), or a SaCas9 nickase (SaCas9n). In some embodiments, the SaCas9 comprises the amino acid sequence SEQ ID NO: 55. In some embodiments, the SaCas9 comprises a N579X mutation of SEQ ID NO: 55, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357, wherein X is any amino acid except for N. In some embodiments, the SaCas9 comprises a N579A mutation of SEQ ID NO: 55, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357.

[00283] In some embodiments, the SaCas9 domain, the SaCas9d domain, or the SaCas9n domain can bind to a nucleic acid sequence having a non-canonical PAM. In some embodiments, the SaCas9 domain, the SaCas9d domain, or the SaCas9n domain can bind to a nucleic acid sequence having a NNGRRT PAM sequence, where N = A, T, C, or G, and R = A or G. In some embodiments, the SaCas9 domain comprises one or more of E781X, N967X, and R1014X mutation of SEQ ID NO: 55, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357, wherein X is any amino acid. In some embodiments, the SaCas9 domain comprises one or more of a E781K, a N967K, and a R1014H mutation of SEQ ID NO: 55, or one or more corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357. In some embodiments, the SaCas9 domain comprises a E781K, a N967K, or a R1014H mutation of SEQ ID NO: 55, or corresponding mutations in any of the amino acid sequences provided in SEQ ID NOs: 108-357.

[00284] In some embodiments, the Cas9 domain of any of the fusion proteins provided herein comprises an amino acid sequence that is 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% identical to any one of SEQ ID NOs: 55-57. In some embodiments, the Cas9 domain of any of the fusion proteins provided herein comprises the amino acid sequence of any one of SEQ ID NOs: 55-57. In some embodiments, the Cas9 domain of any of the fusion proteins provided herein consists of the amino acid sequence of any one of SEQ ID NOs: 55-57.

Exemplary SaCas9 sequence

[00285] KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSK
RGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA
ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGE
VRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPF
GWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLE
YYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDI
TARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLS
LKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFI
QSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKE
NAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL
VKQEE**N**SKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDI
NRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKW
KFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEI
ETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIV
NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKY
YEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRF
DVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNND
LIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKY
STDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 55)

[00286] Residue N579 of SEQ ID NO: 55, which is underlined and in bold, may be mutated (*e.g.*, to a A579) to yield a SaCas9 nickase.

Exemplary SaCas9n sequence

[00287] KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSK RGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGE VRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPF GWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLE YYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDI TARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLS LKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFI QSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKE NAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL VKQEE**A**SKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDI NRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKW KFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEI ETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIV NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKY YEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRF DVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNND LIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKY STDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 56).

[00288] Residue A579 of SEQ ID NO: 56, which can be mutated from N579 of SEQ ID NO: 55 to yield a SaCas9 nickase, is underlined and in bold.

Exemplary SaKKH Cas9

[00289] KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSK RGARRLKRRRRHRIQRVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSA ALLHLAKRRGVHNVNEVEEDTGNELSTKEQISRNSKALEEKYVAELQLERLKKDGE VRGSINRFKTSDYVKEAKQLLKVQKAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPF GWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALNDLNNLVITRDENEKLE YYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNLKVYHDIKDI TARKEIIENAELLDQIAKILTIYQSSEDIQEELTNLNSELTQEEIEQISNLKGYTGTHNLS LKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFI QSIKVINAIIKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEIIRTTGKE NAKYLIEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIPRSVSFDNSFNNKVL VKQEE**A**SKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKKEYLLEERDI

NRFSVQKDFINRNLVDTRYATRGLMNLLRSYFRVNNLDVKVKSINGGFTSFLRRKW KFKKERNKGYKHHAEDALIIANADFIFKEWKKLDKAKKVMENQMFEEKQAESMPEI ETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNR*K*LINDTLYSTRKDDKGNTLIV NNLNGLYDKDNDKLKKLINKSPEKLLMYHHDPQTYQKLKLIMEQYGDEKNPLYKY YEETGNYLTKYSKKDNGPVIKKIKYYGNKLNAHLDITDDYPNSRNKVVKLSLKPYRF DVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFY*K*ND LIKINGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPP*H*IIKTIASKTQSIKKY STDILGNLYEVKSKKHPQIIKKG (SEQ ID NO: 57).

[00290] Residue A579 of SEQ ID NO: 57, which can be mutated from N579 of SEQ ID NO: 55 to yield a SaCas9 nickase, is underlined and in bold. Residues K781, K967, and H1014 of SEQ ID NO: 57, which can be mutated from E781, N967, and R1014 of SEQ ID NO: 55 to yield a SaKKH Cas9 are underlined and in italics.

[00291] In some embodiments, the Cas9 domain is a Cas9 domain from *Streptococcus pyogenes* (SpCas9). In some embodiments, the SpCas9 domain is a nuclease active SpCas9, a nuclease inactive SpCas9 (SpCas9d), or a SpCas9 nickase (SpCas9n). In some embodiments, the SpCas9 comprises the amino acid sequence SEQ ID NO: 58. In some embodiments, the SpCas9 comprises a D9X mutation of SEQ ID NO: 58, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357, wherein X is any amino acid except for D. In some embodiments, the SpCas9 comprises a D9A mutation of SEQ ID NO: 58, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357. In some embodiments, the SpCas9 domain, the SpCas9d domain, or the SpCas9n domain can bind to a nucleic acid sequence having a non-canonical PAM. In some embodiments, the SpCas9 domain, the SpCas9d domain, or the SpCas9n domain can bind to a nucleic acid sequence having a NGG, a NGA, or a NGCG PAM sequence. In some embodiments, the SpCas9 domain comprises one or more of a D1134X, a R1334X, and a T1336X mutation of SEQ ID NO: 58, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-35, wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1134E, R1334Q, and T1336R mutation of SEQ ID NO: 58, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-35. In some embodiments, the SpCas9 domain comprises a D1134E, a R1334Q, and a T1336R mutation of SEQ ID NO: 58, or corresponding mutations in any of the amino acid sequences provided in SEQ ID NOs: 108-35. In some embodiments, the SpCas9 domain comprises one or more of a D1134X, a R1334X, and a T1336X mutation of SEQ ID NO: 58, or a corresponding mutation in any of

the amino acid sequences provided in SEQ ID NOs: 108-35, wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1134V, a R1334Q, and a T1336R mutation of SEQ ID NO: 58, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-35. In some embodiments, the SpCas9 domain comprises a D1134V, a R1334Q, and a T1336R mutation of SEQ ID NO: 58, or corresponding mutations in any of the amino acid sequences provided in SEQ ID NOs: 108-35. In some embodiments, the SpCas9 domain comprises one or more of a D1134X, a G1217X, a R1334X, and a T1336X mutation of SEQ ID NO: 58, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-35, wherein X is any amino acid. In some embodiments, the SpCas9 domain comprises one or more of a D1134V, a G1217R, a R1334Q, and a T1336R mutation of SEQ ID NO: 58, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-35. In some embodiments, the SpCas9 domain comprises a D1134V, a G1217R, a R1334Q, and a T1336R mutation of SEQ ID NO: 58, or corresponding mutations in any of the amino acid sequences provided in SEQ ID NOs: 108-35.

[00292] In some embodiments, the Cas9 domain of any of the fusion proteins provided herein comprises an amino acid sequence that is 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% identical to any one of SEQ ID NOs: 58-62. In some embodiments, the Cas9 domain of any of the fusion proteins provided herein comprises the amino acid sequence of any one of SEQ ID NOs: 58-62. In some embodiments, the Cas9 domain of any of the fusion proteins provided herein consists of the amino acid sequence of any one of SEQ ID NOs: 58-62.

Exemplary SpCas9

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY

NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:
58)

Exemplary SpCas9n

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK

RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO:
59)

Exemplary SpEQR Cas9

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF<u>E</u>SPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA

PAAFKYFDTTIDRK**Q**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 60)

[00293] Residues E1134, Q1334, and R1336 of SEQ ID NO: 60, which can be mutated from D1134, R1334, and T1336 of SEQ ID NO: 58 to yield a SpEQR Cas9, are underlined and in bold.

Exemplary SpVQR Cas9

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRK**Q**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 61)

[00294] Residues V1134, Q1334, and R1336 of SEQ ID NO: 61, which can be mutated from D1134, R1334, and T1336 of SEQ ID NO: 58 to yield a SpVQR Cas9, are underlined and in bold.

Exemplary SpVRER Cas9

DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETA
EATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERH
PIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFRGHFLIEGDL
NPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLENLIAQLPGE
KKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLAQIGDQYADL
FLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQLPEKY
KEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTF
DNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAW
MTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVY
NELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERL
KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRN
FMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVK
VMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQL
QNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDK
NRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIK
RQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDFQFYKV
REINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGK
ATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSM
PQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**SPTVAYSVLVV
AKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFE
LENGRKRMLASA**R**ELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQ
HKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGA
PAAFKYFDTTIDRK**E**Y**R**STKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 62)

[00295] Residues V1134, R1217, Q1334, and R1336 of SEQ ID NO: 62, which can be mutated from D1134, G1217, R1334, and T1336 of SEQ ID NO: 58 to yield a SpVRER Cas9, are underlined and in bold.

High fidelity Cas9 domains

[00296] Some aspects of the disclosure provide high fidelity Cas9 domains of the nucleobase editors provided herein. In some embodiments, high fidelity Cas9 domains are engineered Cas9 domains comprising one or more mutations that decrease electrostatic interactions between the Cas9 domain and the sugar-phosphate backbone of DNA, as compared to a corresponding wild-type Cas9 domain. Without wishing to be bound by any particular theory, high fidelity Cas9 domains that have decreased electrostatic interactions with the sugar-phosphate backbone of DNA may have less off-target effects. In some embodiments, the Cas9 domain (*e.g.,* a wild type Cas9 domain) comprises one or more mutations that decreases the association between the Cas9 domain and the sugar-phosphate backbone of DNA. In some embodiments, a Cas9 domain comprises one or more mutations that decreases the association between the Cas9 domain and the sugar-phosphate backbone of DNA by at least 1%, at least 2%, at least 3%, at least 4%, at least 5%, at least10%, at least 15%, at least 20%, at least 25%, 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%, or more.

[00297] In some embodiments, any of the Cas9 fusion proteins provided herein comprise one or more of N497X, R661X, Q695X, and/or Q926X mutation of the amino acid sequence provided in SEQ ID NO: 52, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357, wherein X is any amino acid. In some embodiments, any of the Cas9 fusion proteins provided herein comprise one or more of N497A, R661A, Q695A, and/or Q926A mutation of the amino acid sequence provided in SEQ ID NO: 52, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357. In some embodiments, the Cas9 domain comprises a D10A mutation of the amino acid sequence provided in SEQ ID NO: 52, or a corresponding mutation in any of the amino acid sequences provided in SEQ ID NOs: 108-357. In some embodiments, the Cas9 domain (*e.g.,* of any of the fusion proteins provided herein) comprises the amino acid sequence as set forth in SEQ ID NO: 62. Cas9 domains with high fidelity are known in the art and would be apparent to the skilled artisan. For example, Cas9 domains with high fidelity have been described in Kleinstiver, B.P., *et al.* "High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects." *Nature* 529, 490-495 (2016); and Slaymaker, I.M., *et al.* "Rationally engineered Cas9 nucleases with improved specificity." *Science* 351, 84-88 (2015); the entire contents of each are incorporated herein by reference.

[00298] It should be appreciated that any of the base editors provided herein, for example, any of the adenosine deaminase base editors provided herein, may be converted into high fidelity base editors by modifying the Cas9 domain as described herein to generate high fidelity base editors, for example, a high fidelity adenosine base editor. In some embodiments, the high fidelity Cas9 domain is a dCas9 domain. In some embodiments, the high fidelity Cas9 domain is a nCas9 domain.

High Fidelity Cas9 domain where mutations relative to Cas9 of SEQ ID NO: 10 are shown in bold and underlines

[00299] DKKYSIGL**A**IGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIG ALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESFLV EEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMIKFR GHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRLE NLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKAL VRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNRE DLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLA RGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMT**A**FDKNLPNEKVLPKHSL LYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYF KKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDR EMIEERLKTYAHLFDDKVMKQLKRRRYTGWG**A**LSRKLINGIRDKQSGKTILDFLKSD GFANRNFM**A**LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKV VDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHP VENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKV LTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELD KAGFIKRQLVETR**A**ITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDFRKDF QFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAKSE QEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVR KVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAY SVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLP KYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQK QLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFT LTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGD (SEQ ID NO: 63)

Nucleic acid programmable DNA binding proteins

[00300] Some aspects of the disclosure provide nucleic acid programmable DNA binding proteins, which may be used to guide a protein, such as a base editor, to a specific nucleic acid (*e.g.,* DNA or RNA) sequence. Nucleic acid programmable DNA binding proteins include, without limitation, Cas9 (*e.g.,* dCas9 and nCas9), CasX, CasY, Cpf1, C2c1, C2c2, C2C3, and Argonaute. One example of an 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 (Cpf1). Similar to Cas9, Cpf1 is also a class 2 CRISPR effector. It has been shown that Cpf1mediates robust DNA interference with features distinct from Cas9. 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.

[00301] Also useful in the present compositions and methods are nuclease-inactive Cpf1 (dCpf1) variants that may be used as a guide nucleotide sequence-programmable DNA-binding protein domain. The 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 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 Cpf1 is responsible for cleaving both DNA strands and inactivation of the RuvC-like domain inactivates Cpf1 nuclease activity. For example, mutations corresponding to D917A, E1006A, or D1255A in *Francisella novicida* Cpf1 (SEQ ID NO: 382) inactivates Cpf1 nuclease activity. In some embodiments, the dCpf1 of the present disclosure comprises mutations corresponding to D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A, E1006A/D1255A, or D917A/ E1006A/D1255A in SEQ ID NO: 376. It is to be understood that any mutations, *e.g.*, substitution mutations, deletions, or insertions that inactivate the RuvC domain of Cpf1, may be used in accordance with the present disclosure.

[00302] In some embodiments, the nucleic acid programmable DNA binding protein

(napDNAbp) of any of the fusion proteins provided herein may be a Cpf1 protein. In some embodiments, the Cpf1 protein is a Cpf1 nickase (nCpf1). In some embodiments, the Cpf1 protein is a nuclease inactive Cpf1 (dCpf1). In some embodiments, the Cpf1, the nCpf1, or the dCpf1 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 ease 99.5% identical to any one of SEQ ID NOs: 376-382. In some embodiments, the dCpf1 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 ease 99.5% identical to any one of SEQ ID NOs: 376-382, and comprises mutations corresponding to D917A, E1006A, D1255A, D917A/E1006A, D917A/D1255A, E1006A/D1255A, or D917A/ E1006A/D1255A in SEQ ID NO: 376. In some embodiments, the dCpf1 comprises an amino acid sequence of any one SEQ ID NOs: 376-382. It should be appreciated that Cpf1 from other bacterial species may also be used in accordance with the present disclosure.

Wild type *Francisella novicida* Cpf1 (SEQ ID NO: 376) (D917, E1006, and D1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**D**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV

VHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**D**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 376)

Francisella novicida Cpf1 D917A (SEQ ID NO: 377) (A917, E1006, and D1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**A**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**D**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 377)

Francisella novicida Cpf1 E1006A (SEQ ID NO: 378) (D917, A1006, and D1255 are bolded

and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**D**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**D**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 378)

Francisella novicida Cpf1 D1255A (SEQ ID NO: 379) (D917, E1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY

SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**D**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**A**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 379)

Francisella novicida Cpf1 D917A/E1006A (SEQ ID NO: 380) (A917, A1006, and D1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR

FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**A**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**D**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 380)

Francisella novicida Cpf1 D917A/D1255A (SEQ ID NO: 381) (A917, E1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**A**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**E**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**A**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 381)

Francisella novicida Cpf1 E1006A/D1255A (SEQ ID NO: 382) (D917, A1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN
TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**D**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**A**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 382)

Francisella novicida Cpf1 D917A/E1006A/D1255A (SEQ ID NO: 383) (A917, A1006, and A1255 are bolded and underlined)

MSIYQEFVNKYSLSKTLRFELIPQGKTLENIKARGLILDDEKRAKDYKKAKQIIDKYH
QFFIEEILSSVCISEDLLQNYSDVYFKLKKSDDDNLQKDFKSAKDTIKKQISEYIKDSE
KFKNLFNQNLIDAKKGQESDLILWLKQSKDNGIELFKANSDITDIDEALEIIKSFKGWT
TYFKGFHENRKNVYSSNDIPTSIIYRIVDDNLPKFLENKAKYESLKDKAPEAINYEQIK
KDLAEELTFDIDYKTSEVNQRVFSLDEVFEIANFNNYLNQSGITKFNTIIGGKFVNGEN

TKRKGINEYINLYSQQINDKTLKKYKMSVLFKQILSDTESKSFVIDKLEDDSDVVTTM
QSFYEQIAAFKTVEEKSIKETLSLLFDDLKAQKLDLSKIYFKNDKSLTDLSQQVFDDY
SVIGTAVLEYITQQIAPKNLDNPSKKEQELIAKKTEKAKYLSLETIKLALEEFNKHRDI
DKQCRFEEILANFAAIPMIFDEIAQNKDNLAQISIKYQNQGKKDLLQASAEDDVKAIK
DLLDQTNNLLHKLKIFHISQSEDKANILDKDEHFYLVFEECYFELANIVPLYNKIRNYI
TQKPYSDEKFKLNFENSTLANGWDKNKEPDNTAILFIKDDKYYLGVMNKKNNKIFD
DKAIKENKGEGYKKIVYKLLPGANKMLPKVFFSAKSIKFYNPSEDILRIRNHSTHTKN
GSPQKGYEKFEFNIEDCRKFIDFYKQSISKHPEWKDFGFRFSDTQRYNSIDEFYREVE
NQGYKLTFENISESYIDSVVNQGKLYLFQIYNKDFSAYSKGRPNLHTLYWKALFDER
NLQDVVYKLNGEAELFYRKQSIPKKITHPAKEAIANKNKDNPKKESVFEYDLIKDKR
FTEDKFFFHCPITINFKSSGANKFNDEINLLLKEKANDVHILSI**A**RGERHLAYYTLVDG
KGNIIKQDTFNIIGNDRMKTNYHDKLAAIEKDRDSARKDWKKINNIKEMKEGYLSQV
VHEIAKLVIEYNAIVVF**A**DLNFGFKRGRFKVEKQVYQKLEKMLIEKLNYLVFKDNEF
DKTGGVLRAYQLTAPFETFKKMGKQTGIIYYVPAGFTSKICPVTGFVNQLYPKYESV
SKSQEFFSKFDKICYNLDKGYFEFSFDYKNFGDKAAKGKWTIASFGSRLINFRNSDKN
HNWDTREVYPTKELEKLLKDYSIEYGHGECIKAAICGESDKKFFAKLTSVLNTILQM
RNSKTGTELDYLISPVADVNGNFFDSRQAPKNMPQDA**A**ANGAYHIGLKGLMLLGRI
KNNQEGKKLNLVIKNEEYFEFVQNRNN (SEQ ID NO: 383)

[00303] 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.*, 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. The sequence of *Natronobacterium gregoryi* Argonaute is provided in SEQ ID NO: 416.

Wild type *Natronobacterium gregoryi* Argonaute (SEQ ID NO: 416)
MTVIDLDSTTTADELTSGHTYDISVTLTGVYDNTDEQHPRMSLAFEQDNGERRYITL
WKNTTPKDVFTYDYATGSTYIFTNIDYEVKDGYENLTATYQTTVENATAQEVGTTD
EDETFAGGEPLDHHLDDALNETPDDAETESDSGHVMTSFASRDQLPEWTLHTYTLT
ATDGAKTDTEYARRTLAYTVRQELYTDHDAAPVATDGLMLLTPEPLGETPLDLDCG
VRVEADETRTLDYTTAKDRLLARELVEEGLKRSLWDDYLVRGIDEVLSKEPVLTCD
EFDLHERYDLSVEVGHSGRAYLHINFRHRFVPKLTLADIDDDNIYPGLRVKTTYRPR
RGHIVWGLRDECATDSLNTLGNQSVVAYHRNNQTPINTDLLDAIEAADRRVVETRR
QGHGDDAVSFPQELLAVEPNTHQIKQFASDGFHQQARSKTRLSASRCSEKAQAFAER
LDPVRLNGSTVEFSSEFFTGNNEQQLRLLYENGESVLTFRDGARGAHPDETFSKGIVN
PPESFEVAVVLPEQQADTCKAQWDTMADLLNQAGAPPTRSETVQYDAFSSPESISLN
VAGAIDPSEVDAAFVVLPPDQEGFADLASPTETYDELKKALANMGIYSQMAYFDRF
RDAKIFYTRNVALGLLAAAGGVAFTTEHAMPGDADMFIGIDVSRSYPEDGASGQINI
AATATAVYKDGTILGHSSTRPQLGEKLQSTDVRDIMKNAILGYQQVTGESPTHIVIHR
DGFMNEDLDPATEFLNEQGVEYDIVEIRKQPQTRLLAVSDVQYDTPVKSIAAINQNEP
RATVATFGAPEYLATRDGGGLPRPIQIERVAGETDIETLTRQVYLLSQSHIQVHNSTA
RLPITTAYADQASTHATKGYLVQTGAFESNVGFL (SEQ ID NO: 416)

[00304] 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 Argunaute (MpAgo) protein. The CRISPR-associated Marinitoga piezophila Argunaute (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.

[00305] In some embodiments, the nucleic acid programmable DNA binding protein (napDNAbp) is a single effector of a microbial CRISPR-Cas system. Single effectors of microbial CRISPR-Cas systems include, without limitation, Cas9, Cpf1, C2c1, C2c2, and C2c3. Typically, microbial CRISPR-Cas systems are divided into Class 1 and Class 2 systems. Class 1 systems have multisubunit effector complexes, while Class 2 systems have a single protein effector. For example, Cas9 and Cpf1 are Class 2 effectors. In addition to Cas9 and Cpf1, three distinct Class 2 CRISPR-Cas systems (C2c1, C2c2, and C2c3) 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 is hereby incorporated by reference. Effectors of two of the systems, C2c1 and C2c3, contain RuvC-like endonuclease domains related to Cpf1. A third system, C2c2 contains an effector with two predicated HEPN RNase domains. Production of mature CRISPR RNA is tracrRNA-independent, unlike production of CRISPR RNA by C2c1. C2c1 depends on both CRISPR RNA and tracrRNA for DNA cleavage. Bacterial C2c2 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 Cpf1. See, e.g., East-Seletsky, et al., "Two distinct RNase activities of CRISPR-C2c2 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 C2c2 in *Leptotrichia shahii* has shown that C2c2 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.

[00306] The crystal structure of *Alicyclobaccillus acidoterrastris* C2c1 (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.

[00307] In some embodiments, the nucleic acid programmable DNA binding protein (napDNAbp) of any of the fusion proteins provided herein 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 C2c2 protein. In some embodiments, the napDNAbp is a C2c3 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 ease 99.5% identical to a naturally-occurring C2c1, C2c2, or C2c3 protein. In some embodiments, the napDNAbp is a naturally-occurring C2c1, C2c2, or C2c3 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 ease 99.5% identical to any one of SEQ ID NOs: 438 or 439. In some embodiments, the napDNAbp comprises an amino acid sequence of any one SEQ ID NOs: 438 or 439. It should be appreciated that C2c1, C2c2, or C2c3 from other bacterial species may also be used in accordance with the present disclosure.

C2c1 (uniprot.org/uniprot/T0D7A2#)

sp|T0D7A2|C2C1_ALIAG CRISPR-associated endonuclease C2c1 OS=Alicyclobacillus acidoterrestris (strain ATCC 49025 / DSM 3922 / CIP 106132 / NCIMB 13137 / GD3B) GN=c2c1 PE=1 SV=1

MAVKSIKVKLRLDDMPEIRAGLWKLHKEVNAGVRYYTEWLSLLRQENLYRRSPNG
DGEQECDKTAEECKAELLERLRARQVENGHRGPAGSDDELLQLARQLYELLVPQAI
GAKGDAQQIARKFLSPLADKDAVGGLGIAKAGNKPRWVRMREAGEPGWEEEKEKA
ETRKSADRTADVLRALADFGLKPLMRVYTDSEMSSVEWKPLRKGQAVRTWDRDM
FQQAIERMMSWESWNQRVGQEYAKLVEQKNRFEQKNFVGQEHLVHLVNQLQQDM
KEASPGLESKEQTAHYVTGRALRGSDKVFEKWGKLAPDAPFDLYDAEIKNVQRRNT
RRFGSHDLFAKLAEPEYQALWREDASFLTRYAVYNSILRKLNHAKMFATFTLPDAT
AHPIWTRFDKLGGNLHQYTFLFNEFGERRHAIRFHKLLKVENGVAREVDDVTVPISM

SEQLDNLLPRDPNEPIALYFRDYGAEQHFTGEFGGAKIQCRRDQLAHMHRRRGARD
VYLNVSVRVQSQSEARGERRPPYAAVFRLVGDNHRAFVHFDKLSDYLAEHPDDGKL
GSEGLLSGLRVMSVDLGLRTSASISVFRVARKDELKPNSKGRVPFFFPIKGNDNLVAV
HERSQLLKLPGETESKDLRAIREERQRTLRQLRTQLAYLRLLVRCGSEDVGRRERSW
AKLIEQPVDAANHMTPDWREAFENELQKLKSLHGICSDKEWMDAVYESVRRVWRH
MGKQVRDWRKDVRSGERPKIRGYAKDVVGGNSIEQIEYLERQYKFLKSWSFFGKVS
GQVIRAEKGSRFAITLREHIDHAKEDRLKKLADRIIMEALGYVYALDERGKGKWVA
KYPPCQLILLEELSEYQFNNDRPPSENNQLMQWSHRGVFQELINQAQVHDLLVGTM
YAAFSSRFDARTGAPGIRCRRVPARCTQEHNPEPFPWWLNKFVVEHTLDACPLRAD
DLIPTGEGEIFVSPFSAEEGDFHQIHADLNAAQNLQQRLWSDFDISQIRLRCDWGEVD
GELVLIPRLTGKRTADSYSNKVFYTNTGVTYYERERGKKRRKVFAQEKLSEEEAELL
VEADEAREKSVVLMRDPSGIINRGNWTRQKEFWSMVNQRIEGYLVKQIRSRVPLQD
SACENTGDI (SEQ ID NO: 438)

C2c2 (uniprot.org/uniprot/P0DOC6)
>sp|P0DOC6|C2C2_LEPSD CRISPR-associated endoribonuclease C2c2 OS=Leptotrichia shahii (strain DSM 19757 / CCUG 47503 / CIP 107916 / JCM 16776 / LB37) GN=c2c2 PE=1 SV=1
MGNLFGHKRWYEVRDKKDFKIKRKVKVKRNYDGNKYILNINENNNKEKIDNNKFIR
KYINYKKNDNILKEFTRKFHAGNILFKLKGKEGIIRIENNDDFLETEEVVLYIEAYGKS
EKLKALGITKKKIIDEAIRQGITKDDKKIEIKRQENEEEIEIDIRDEYTNKTLNDCSIILRI
IENDELETKKSIYEIFKNINMSLYKIIEKIIENETEKVFENRYYEEHLREKLLKDDKIDVI
LTNFMEIREKIKSNLEILGFVKFYLNVGGDKKKSKNKKMLVEKILNINVDLTVEDIAD
FVIKELEFWNITKRIEKVKKVNNEFLEKRRNRTYIKSYVLLDKHEKFKIERENKKDKI
VKFFVENIKNNSIKEKIEKILAEFKIDELIKKLEKELKKGNCDTEIFGIFKKHYKVNFDS
KKFSKKSDEEKELYKIIYRYLKGRIEKILVNEQKVRLKKMEKIEIEKILNESILSEKILK
RVKQYTLEHIMYLGKLRHNDIDMTTVNTDDFSRLHAKEELDLELITFFASTNMELNK
IFSRENINNDENIDFFGGDREKNYVLDKKILNSKIKIIRDLDFIDNKNNITNNFIRKFTKI
GTNERNRILHAISKERDLQGTQDDYNKVINIIQNLKISDEEVSKALNLDVVFKDKKNII
TKINDIKISEENNNDIKYLPSFSKVLPEILNLYRNNPKNEPFDTIETEKIVLNALIYVNKE
LYKKLILEDDLEENESKNIFLQELKKTLGNIDEIDENIIENYYKNAQISASKGNNKAIK
KYQKKVIECYIGYLRKNYEELFDFSDFKMNIQEIKKQIKDINDNKTYERITVKTSDKTI
VINDDFEYIISIFALLNSNAVINKIRNRFFATSVWLNTSEYQNIIDILDEIMQLNTLRNEC
ITENWNLNLEEFIQKMKEIEKDFDDFKIQTKKEIFNNYYEDIKNNILTEFKDDINGCDV

LEKKLEKIVIFDDETKFEIDKKSNILQDEQRKLSNINKKDLKKKVDQYIKDKDQEIKS
KILCRIIFNSDFLKKYKKEIDNLIEDMESENENKFQEIYYPKERKNELYIYKKNLFLNIG
NPNFDKIYGLISNDIKMADAKFLFNIDGKNIRKNKISEIDAILKNLNDKLNGYSKEYKE
KYIKKLKENDDFFAKNIQNKNYKSFEKDYNRVSEYKKIRDLVEFNYLNKIESYLIDIN
WKLAIQMARFERDMHYIVNGLRELGIIKLSGYNTGISRAYPKRNGSDGFYTTTAYYK
FFDEESYKKFEKICYGFGIDLSENSEINKPENESIRNYISHFYIVRNPFADYSIAEQIDRV
SNLLSYSTRYNNSTYASVFEVFKKDVNLDYDELKKKFKLIGNNDILERLMKPKKVSV
LELESYNSDYIKNLIIELLTKIENTNDTL (SEQ ID NO: 439)

Fusion proteins comprising a nuclease programmable DNA binding protein and an adenosine deaminase

[00308] Some aspects of the disclosure provide fusion proteins comprising a nucleic acid programmable DNA binding protein (napDNAbp) and an adenosine deaminase. In some embodiments, any of the fusion proteins provided herein are base editors. In some embodiments, the napDNAbp is a Cas9 domain, a Cpf1 domain, a CasX domain, a CasY domain, a C2c1 domain, a C2c2 domain, aC2c3 domain, or an Argonaute domain. In some embodiments, the napDNAbp is any napDNAbp provided herein. Some aspects of the disclosure provide fusion proteins comprising a Cas9 domain and an adenosine deaminase. The Cas9 domain may be any of the Cas9 domains or Cas9 proteins (*e.g.*, dCas9 or nCas9) provided herein. In some embodiments, any of the Cas9 domains or Cas9 proteins (*e.g.*, dCas9 or nCas9) provided herein may be fused with any of the adenosine deaminases provided herein. In some embodiments, the fusion protein comprises the structure:

NH_2-[adenosine deaminase]-[napDNAbp]-COOH; or
NH_2-[napDNAbp]-[adenosine deaminase]-COOH

[00309] In some embodiments, the fusion proteins comprising an adenosine deaminase and a napDNAbp (*e.g.,* Cas9 domain) do not include a linker sequence. In some embodiments, a linker is present between the adenosine deaminase domain and the napDNAbp . In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker. In some embodiments, the adenosine deaminase and the napDNAbp are fused via any of the linkers provided herein. For example, in some embodiments the adenosine deaminase and the napDNAbp are fused via any of the linkers provided below in the section entitled "Linkers". In some embodiments, the adenosine

deaminase and the napDNAbp are fused via a linker that comprises between 1 and and 200 amino acids. In some embodiments, the adenosine deaminase and the napDNAbp are fused via a linker that comprises from 1 to 5, 1 to 10, 1 to 20, 1 to 30, 1 to 40, 1 to 50, 1 to 60, 1 to 80, 1 to 100, 1 to 150, 1 to 200, 5 to 10, 5 to 20, 5 to 30, 5 to 40, 5 to 60, 5 to 80, 5 to 100, 5 to 150, 5 to 200, 10 to 20, 10 to 30, 10 to 40, 10 to 50, 10 to 60, 10 to 80, 10 to 100, 10 to 150, 10 to 200, 20 to 30, 20 to 40, 20 to 50, 20 to 60, 20 to 80, 20 to 100, 20 to 150, 20 to 200, 30 to 40, 30 to 50, 30 to 60, 30 to 80, 30 to 100, 30 to 150, 30 to 200, 40 to 50, 40 to 60, 40 to 80, 40 to 100, 40 to 150, 40 to 200, 50 to 60 50 to 80, 50 to 100, 50 to 150, 50 to 200, 60 to 80, 60 to 100, 60 to 150, 60 to 200, 80 to 100, 80 to 150, 80 to 200, 100 to 150, 100 to 200, or 150 to 200 amino acids in length. In some embodiments, the adenosine deaminase and the napDNAbp are fused via a linker that comprises 4, 16, 32, or 104 amino acids in length. In some embodiments, the adenosine deaminase and the napDNAbp are fused via a linker that comprises the amino acid sequence of SGSETPGTSESATPES (SEQ ID NO: 10), SGGS (SEQ ID NO: 37), SGGSSGSETPGTSESATPESSGGS (SEQ ID NO: 384), SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 385), or GGSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSGGS (SEQ ID NO: 386). In some embodiments, the adenosine deaminase and the napDNAbp are fused via a linker comprising the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 10), which may also be referred to as the XTEN linker. In some embodiments, the linker is 24 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPES (SEQ ID NO: 685). In some embodiments, the linker is 40 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS (SEQ ID NO: 686). In some embodiments, the linker is 64 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 687). In some embodiments, the linker is 92 amino acids in length. In some embodiments, the linker comprises the amino acid sequence PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATS (SEQ ID NO: 688).

Fusion proteins comprising an inhibitor of base repair

[00310] Some aspects of the disclosure provide fusion proteins that comprise an

inhibitor of base repair (IBR). For example a fusion protein comprising an adenosine deaminase and a nucleic acid programmable DNA binding protein may further comprise an inhibitor of base repair. In some embodiments, the IBR comprises an inhibitor of inosine base repair. In some embodiments, the IBR is an inhibitor of inosine base excision repair. In some embodiments, the inhibitor of inosine base excision repair is a catalytically inactive inosine specific nuclease (dISN).

[00311] In some embodiments, the fusion proteins provided herein further comprise a catalytically inactive inosine-specific nuclease (dISN). In some embodiments, any of the fusion proteins provided herein that comprise a napDNAbp (*e.g.,* a nuclease active Cas9 domain, a nuclease inactive dCas9 domain, or a Cas9 nickase) and an adenosine deaminase may be further fused to a catalytically inactive inosine-specific nuclease (dISN) either directly or via a linker. Some aspects of this disclosure provide fusion proteins that comprise an adenosine deaminase (*e.g.,* an engineered adenosine deaminase that deaminates adenosine in a DNA) a napDNAbp (*e.g.,* a dCas9 or nCas9), and a dISN. Without wishing to be bound by any particular theory, cellular DNA-repair response to the presence of I:T heteroduplex DNA may be responsible for the decrease in nucleobase editing efficiency in cells. For example, AAG catalyzes removal of inosine (I) from DNA in cells, which may initiate base excision repair, with reversion of the I:T pair to a A:T pair as the most common outcome. In some embodiments, a catalytically inactive inosine-specific nuclease may be capable of binding an inosine in a nucleic acid, without cleaving the nucleic acid, to prevent removal (*e.g.,* by cellular DNA repair mechanisms) of the inosine residue in the DNA.

[00312] In some embodiments, a dISN may inhibit (*e.g.,* by steric hindrance) inosine removing enzymes from excising the inosine residue from DNA. For example, catalytically dead inosine glycosylases (e.g., alkyl adenine glycosylase [AAG]) will bind inosine but will not create an abasic site or remove the inosine, thereby sterically blocking the newly-formed inosine moiety from potential DNA damage/repair mechanisms. Thus, this disclosure contemplates a fusion protein comprising a napDNAbp and an adenosine deaminase further fused to a dISN . This disclosure contemplates a fusion protein comprising any Cas9 domain, for example, a Cas9 nickase (nCas9) domain, a catalytically inactive Cas9 (dCas9) domain, a high fidelity Cas9 domain, or a Cas9 domain with reduced PAM exclusivity. It should be understood that the use of a dISN may increase the editing efficiency of a adenosine deaminase that is capable of catalyzing a A to I change. For example, fusion proteins comprising a dISN domain may be more efficient in deaminating A residues. In some embodiments, the fusion protein comprises the structure:

NH_2-[adenosine deaminase]-[napDNAbp]-[dISN]-COOH;
NH_2-[adenosine deaminase]-[dISN]-[napDNAbp]-COOH;
NH_2-[dISN] -[adenosine deaminase]-[napDNAbp]-COOH;
NH_2-[napDNAbp]-[adenosine deaminase]-[dISN]-COOH;
NH_2-[napDNAbp]-[dISN]-[adenosine deaminase]-COOH; or
NH_2-[dISN]-[napDNAbp]-[adenosine deaminase]-COOH

[00313] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between two domains or proteins (*e.g.*, adenosine deaminase, napDNAbp, or dISN). In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker sequence. In some embodiments, a dISN comprises an inosine-specific nuclease that has reduced or nuclease activity, or does not have nuclease activity. In some embodiments, a dISN has up to 1%, up to 2%, up to 3%, up to 4%, up to 5%, up to 10%, up to 15%, up to 20%, up to 25%, up to 30%, up to 35%, up to 40%, up to 45%, or up to 50% of the nuclease activity of a corresponding (*e.g.*, the wild-type) inosine-specific nuclease. In some embodiments, the dISN is a wild-type inosine-specific nuclease that comprises one or more mutations that reduces or eliminates the nuclease activity of the wild-type inosine-specific nuclease. Exemplary catalytically inactive inosine-specific nucleases include, without limitation, catalytically inactive AAG nuclease and catalytically inactive EndoV nuclease. In some embodiments, the catalytically inactive AAG nuclease comprises an E125Q mutation as compared to SEQ ID NO: 32, or a corresponding mutation in another AAG nuclease. In some embodiments, the catalytically inactive AAG nuclease comprises the amino acid sequence set forth in SEQ ID NO: 32. In some embodiments, the catalytically inactive EndoV nuclease comprises an D35A mutation as compared to SEQ ID NO 32, or a corresponding mutation in another EndoV nuclease. In some embodiments, the catalytically inactive EndoV nuclease comprises the amino acid sequence set forth in SEQ ID NO: 33. It should be appreciated that other catalytically inactive inosine-specific nucleases (dISNs) would be apparent to the skilled artisan and are within the scope of this disclosure.

[00314] In some embodiments, the dISN proteins provided herein include fragments of dISN proteins and proteins homologous to a dISN or a dISN fragment. For example, in some embodiments, a dISN comprises a fragment of the amino acid sequence set forth in SEQ ID NO: 32 or 33. In some embodiments, a dISN 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: 32 or 33. In some embodiments, a dISN comprises an amino acid sequence homologous to the amino acid sequence set forth in SEQ ID NO: 32 or 33, or an amino acid sequence homologous to a fragment of the amino acid sequence set forth in SEQ ID NO: 32 or 33. In some embodiments, proteins comprising a dISN or fragments of a dISN or homologs of a dISN or a dISN fragment are referred to as "dISN variants." A dISN variant shares homology to a dISN, or a fragment thereof. For example a dISN 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 dISN or a dISN as set forth in SEQ ID NO: 32 or 33. In some embodiments, the dISN variant comprises a fragment of dISN, 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 dISN or a dISN as set forth in SEQ ID NO: 32 or 33. In some embodiments, the dISN comprises the following amino acid sequence:

AAG nuclease (E125Q); mutated residue underlined in bold.
KGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELRGRIVET**Q**AYLGPEDEAAH SRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDGACVLLRALEPLEGLET MRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRDLAQDEAVWLERGP LEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDRVAEQDTQA (SEQ ID NO: 32)

EndoV nuclease (D35A); mutated residue underlined in bold.
DLASLRAQQIELASSVIREDRLDKDPPDLIAGA**A**VGFEQGGEVTRAAMVLLKYPSLE LVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVFVDGHGISHPRRLGV ASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQLAWVWRSKARCNP LFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPAFVRYTANQP (SEQ ID NO: 33)

[00315] Suitable dISN proteins are provided herein and additional suitable dISN proteins are known to those in the art, and include, for example, AAG, EndoV, and variants thereof. It should be appreciated that additional proteins that block or inhibit base-excision repair, such as base excision of an inosine, are also within the scope of this disclosure. In some embodiments, a protein that binds inosine in DNA is used.

[00316] Some aspects of the disclosure relate to fusion proteins that comprise MBD4, or TDG, which may be used as inhibitors of base repair. Thus, this disclosure contemplates a fusion protein comprising a napDNAbp and an adenosine deaminase further fused to MBD4 or TDG. This disclosure contemplates a fusion protein comprising any Cas9 domain, for example, a Cas9 nickase (nCas9) domain, a catalytically inactive Cas9 (dCas9) domain, a high fidelity Cas9 domain, or a Cas9 domain with reduced PAM exclusivity. It should be understood that the use of MBD4 or TDG may increase the editing efficiency of a adenosine deaminase that is capable of catalyzing a A to I change. For example, fusion proteins comprising MBD4 or TDG may be more efficient in deaminating A residues. In some embodiments, the fusion protein comprises the structure:

NH_2-[adenosine deaminase]-[napDNAbp]-[MBD4 or TDG]-COOH;
NH_2-[adenosine deaminase]-[MBD4 or TDG]-[napDNAbp]-COOH;
NH_2-[MBD4 or TDG] -[adenosine deaminase]-[napDNAbp]-COOH;
NH_2-[napDNAbp]-[adenosine deaminase]-[MBD4 or TDG]-COOH;
NH_2-[napDNAbp]-[MBD4 or TDG]-[adenosine deaminase]-COOH; or
NH_2-[MBD4 or TDG]-[napDNAbp]-[adenosine deaminase]-COOH

[00317] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between two domains or proteins (*e.g.*, adenosine deaminase, napDNAbp, MBD4 or TDG). In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker sequence. In some embodiments, the MBD4 or TDG is a wild-type MBD4 or TDG. Exemplary, MBD4 and TDG amino acid sequences would be apparent to the skilled artisan and include, without limitation, the MBD4 and TDG amino acid sequences provided below.

Sequence of MBD4:
GTTGLESLSLGDRGAAPTVTSSERLVPDPPNDLRKEDVAMELERVGEDEEQMMIKRS
SECNPLLQEPIASAQFGATAGTECRKSVPCGWERVVKQRLFGKTAGRFDVYFISPQG
LKFRSKSSLANYLHKNGETSLKPEDFDFTVLSKRGIKSRYKDCSMAALTSHLQNQSN
NSNWNLRTRSKCKKDVFMPPSSSSELQESRGLSNFTSTHLLLKEDEGVDDVNFRKVR
KPKGKVTILKGIPIKKTKKGCRKSCSGFVQSDSKRESVCNKADAESEPVAQKSQLDR
TVCISDAGACGETLSVTSEENSLVKKKERSLSSGSNFCSEQKTSGIINKFCSAKDSEHN
EKYEDTFLESEEIGTKVEVVERKEHLHTDILKRGSEMDNNCSPTRKDFTGEKIFQEDTI
PRTQIERRKTSLYFSSKYNKEALSPPRRKAFKKWTPPRSPFNLVQETLFHDPWKLLIA
TIFLNRTSGKMAIPVLWKFLEKYPSAEVARTADWRDVSELLKPLGLYDLRAKTIVKF
SDEYLTKQWKYPIELHGIGKYGNDSYRIFCVNEWKQVHPEDHKLNKYHDWLWENH
EKLSLS (SEQ ID NO: 689)

Sequence of TDG:
EAENAGSYSLQQAQAFYTFPFQQLMAEAPNMAVVNEQQMPEEVPAPAPAQEPVQE
APKGRKRKPRTTEPKQPVEPKKPVESKKSGKSAKSKEKQEKITDTFKVKRKVDRFNG
VSEAELLTKTLPDILTFNLDIVIIGINPGLMAAYKGHHYPGPGNHFWKCLFMSGLSEV
QLNHMDDHTLPGKYGIGFTNMVERTTPGSKDLSSKEFREGGRILVQKLQKYQPRIAV
FNGKCIYEIFSKEVFGVKVKNLEFGLQPHKIPDTETLCYVMPSSSARCAQFPRAQDKV
HYYIKLKDLRDQLKGIERNMDVQEVQYTFDLQLAQEDAKKMAVKEEKYDPGYEAA
YGGAYGENPCSSEPCGFSSNGLIESVELRGESAFSGIPNGQWMTQSFTDQIPSFSNHC
GTQEQEEESHA (SEQ ID NO: 690)

[00318] In some embodiments, the MBD4 or TDG proteins provided herein include fragments of MBD4 or TDG proteins and proteins homologous to a MBD4 or a TDG fragment. For example, in some embodiments, a MBD4 or TDG protein comprises a fragment of the amino acid sequence set forth in SEQ ID NO: 689 or 690. In some embodiments, a MBD4 or TDG 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: 689 or 690. In some embodiments, a MBD4 or TDG protein comprises an amino acid sequence homologous to the amino acid sequence set forth in SEQ ID NO: 689 or 690, or an amino acid sequence homologous to a fragment of the amino acid sequence set forth in SEQ ID NO: 689 or 690. In some embodiments, proteins comprising a MBD4 or TDG or fragments of a MBD4 or TDG or homologs of a MBD4 or TDG fragment are referred to as "MBD4 varients" or "TDG variants." A MBD4 or TDG variant shares homology to a MBD4 or TDG, or a fragment thereof. For example a MBD4 or TDG 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 MBD4 or TDG or a MBD4 or TDG as set forth in SEQ

ID NO: 689 or 690. In some embodiments, the MBD4 or TDG variant comprises a fragment of MBD4 or TDG, 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 MBD4 or TDG or a MBD4 or TDG as set forth in SEQ ID NO: 689 or 690. In some embodiments, the dISN comprises the following amino acid sequence:

[00319] Some aspects of the disclosure relate to fusion proteins that comprise a uracil glycosylase inhibitor (UGI) domain. In some embodiments, any of the fusion proteins provided herein that comprise a napDNAbp (*e.g.,* a nuclease active Cas9 domain, a nuclease inactive dCas9 domain, or a Cas9 nickase), and an adenosine deaminase, may be further fused to a UGI domain either directly or via a linker. Some aspects of this disclosure provide fusion proteins that comprise an adenosine deaminase (*e.g.,* an engineered adenosine deaminase that deaminates deoxyadenosine in a DNA) a napDNAbp (*e.g.,* a dCas9 or nCas9), and a UGI domain. Without wishing to be bound by any particular theory, the cellular DNA-repair response to the presence of I:T heteroduplex DNA may be responsible for the decrease in nucleobase editing efficiency in cells. For example, alkyl adenosine glycosylase (AAG) is involved in inosine (I) associated DNA repair and catalyzes removal of I from DNA in cells. This may initiate base excision repair, with reversion of the I:T pair to a A:T pair as the most common outcome. A UGI domain, may inhibit (*e.g.,* by steric hindrance) inosine removing enzymes from excising the inosine residue from DNA. Thus, this disclosure contemplates a fusion protein comprising a Cas9 domain and an adenosine deaminase domain further fused to a UGI domain. This disclosure contemplates a fusion protein comprising any nucleic acid programmable DNA binding protein, for example, a Cas9 nickase (nCas9) domain, a catalytically inactive Cas9 (dCas9) domain, a high fidelity Cas9 domain, or a Cas9 domain with reduced PAM exclusivity. It should be understood that the use of a UGI domain may increase the editing efficiency of a adenosine deaminase that is capable of catalyzing a A to I change. For example, fusion proteins comprising a UGI domain may be more efficient in deaminating adenosine residues. In some embodiments, the fusion protein comprises the structure:

NH_2-[adenosine deaminase]-[napDNAbp]-[UGI]-COOH;
NH_2-[adenosine deaminase]-[UGI]-[napDNAbp]-COOH;
NH_2-[UGI]-[adenosine deaminase]-[napDNAbp]-COOH;

NH_2-[napDNAbp]-[adenosine deaminase]-[UGI]-COOH;
NH_2-[napDNAbp]-[UGI]-[adenosine deaminase]-COOH; or
NH_2-[UGI]-[napDNAbp]-[adenosine deaminase]-COOH

[00320] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between any of the domains or proteins (*e.g.,* adenosine deaminase, napDNAbp, and/or UGI domains). In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker.

[00321] In some embodiments, a UGI domain comprises a wild-type UGI or a UGI as set forth in SEQ ID NO: 3. 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: 3. 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: 3. In some embodiments, a UGI comprises an amino acid sequence homologous to the amino acid sequence set forth in SEQ ID NO: 3 or an amino acid sequence homologous to a fragment of the amino acid sequence set forth in SEQ ID NO: 3. 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: 3. 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: 3. In some embodiments, the UGI comprises the following amino acid sequence:

>sp|P14739|UNGI_BPPB2 Uracil-DNA glycosylase inhibitor
MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHTAYDESTDENVMLLT
SDAPEYKPWALVIQDSNGENKIKML (SEQ ID NO: 3)

[00322] Suitable UGI protein and nucleotide sequences are provided herein and additional suitable UGI sequences are known to those in the art, and include, for example, those published in Wang *et al.*, Uracil-DNA glycosylase inhibitor gene of bacteriophage PBS2 encodes a binding protein specific for uracil-DNA glycosylase. *J. Biol. Chem.* 264:1163-1171(1989); Lundquist et al., Site-directed mutagenesis and characterization of uracil-DNA glycosylase inhibitor protein. Role of specific carboxylic amino acids in complex formation with *Escherichia coli* uracil-DNA glycosylase. *J. Biol. Chem.* 272:21408-21419(1997); Ravishankar *et al.*, X-ray analysis of a complex of *Escherichia coli* uracil DNA glycosylase (EcUDG) with a proteinaceous inhibitor. The structure elucidation of a prokaryotic UDG. *Nucleic Acids Res.* 26:4880-4887(1998); and Putnam *et al.*, Protein mimicry of DNA from crystal structures of the uracil-DNA glycosylase inhibitor protein and its complex with Escherichia coli uracil-DNA glycosylase. *J. Mol. Biol.* 287:331-346(1999), the entire contents of each are incorporated herein by reference.

[00323] It should be appreciated that additional proteins that block or inhibit base-excision repair, such as base excision of an inosine, are also within the scope of this disclosure. In some embodiments, a protein that binds DNA is used. In another embodiment, a substitute for UGI is used. In some embodiments, a uracil glycosylase inhibitor is a protein that binds single-stranded DNA. For example, a uracil glycosylase inhibitor may be a *Erwinia tasmaniensis* single-stranded binding protein. In some embodiments, the single-stranded binding protein comprises the amino acid sequence (SEQ ID NO: 29). In some embodiments, a uracil glycosylase inhibitor is a protein that binds uracil. In some embodiments, a uracil glycosylase inhibitor is a protein that binds uracil in DNA. In some embodiments, a uracil glycosylase inhibitor is a catalytically inactive uracil DNA-glycosylase protein. In some embodiments, a uracil glycosylase inhibitor is a catalytically inactive uracil DNA-glycosylase protein that does not excise uracil from the DNA. For example, a uracil glycosylase inhibitor is a UdgX. In some embodiments, the UdgX comprises the amino acid sequence (SEQ ID NO: 30). As another example, a uracil glycosylase inhibitor is a catalytically inactive UDG. In some embodiments, a catalytically inactive UDG comprises the amino acid sequence (SEQ ID NO: 31). It should be appreciated that other uracil glycosylase inhibitors would be apparent to the skilled artisan and are within the scope of this

disclosure. In some embodiments, a uracil glycosylase inhibitor is a protein that is homologous to any one of SEQ ID NOs: 29-31. In some embodiments, a uracil glycosylase inhibitor is a protein that is at least 50% identical, at least 55% identical, at least 60% identical, at least 65% identical, 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 98% identical, at least 99% identical, or at least 99.5% identical to any one of SEQ ID NOs: 29-31.

Erwinia tasmaniensis SSB (themostable single-stranded DNA binding protein)
MASRGVNKVILVGNLGQDPEVRYMPNGGAVANITLATSESWRDKQTGETKEKTEW
HRVVLFGKLAEVAGEYLRKGSQVYIEGALQTRKWTDQAGVEKYTTEVVVNVGGT
MQMLGGRSQGGGASAGGQNGGSNNGWGQPQQPQGGNQFSGGAQQQARPQQQPQ
QNNAPANNEPPIDFDDDIP (SEQ ID NO: 29)

UdgX (binds to Uracil in DNA but does not excise)
MAGAQDFVPHTADLAELAAAAGECRGCGLYRDATQAVFGAGGRSARIMMIGEQPG
DKEDLAGLPFVGPAGRLLDRALEAADIDRDALYVTNAVKHFKFTRAAGGKRRIHKT
PSRTEVVACRPWLIAEMTSVEPDVVVLLGATAAKALLGNDFRVTQHRGEVLHVDDV
PGDPALVATVHPSSLLRGPKEERESAFAGLVDDLRVAADVRP (SEQ ID NO: 30)

UDG (catalytically inactive human UDG, binds to Uracil in DNA but does not excise)
MIGQKTLYSFFSPSPARKRHAPSPEPAVQGTGVAGVPEESGDAAAIPAKKAPAGQEE
PGTPPSSPLSAEQLDRIQRNKAAALLRLAARNVPVGFGESWKKHLSGEFGKPYFIKL
MGFVAEERKHYTVYPPPHQVFTWTQMCDIKDVKVVILGQEPYHGPNQAHGLCFSV
QRPVPPPPSLENIYKELSTDIEDFVHPGHGDLSGWAKQGVLLLNAVLTVRAHQANSH
KERGWEQFTDAVVSWLNQNSNGLVFLLWGSYAQKKGSAIDRKRHHVLQTAHPSPL
SVYRGFFGCRHFSKTNELLQKSGKKPIDWKEL (SEQ ID NO: 31)

Fusion proteins comprising a nuclear localization sequence (NLS)

[00324] In some embodiments, the fusion proteins provided herein further comprise one or more nuclear targeting sequences, for example, a nuclear localization sequence (NLS). In some embodiments, a NLS comprises an amino acid sequence that facilitates the importation of a protein, that comprises an NLS, into the cell nucleus (*e.g.*, by nuclear transport). In some embodiments, any of the fusion proteins provided herein further

comprise a nuclear localization sequence (NLS). In some embodiments, the NLS is fused to the N-terminus of the fusion protein. In some embodiments, the NLS is fused to the C-terminus of the fusion protein. In some embodiments, the NLS is fused to the N-terminus of the IBR (*e.g.*, dISN). In some embodiments, the NLS is fused to the C-terminus of the IBR (*e.g.,* dISN). In some embodiments, the NLS is fused to the N-terminus of the napDNAbp. In some embodiments, the NLS is fused to the C-terminus of the napDNAbp. In some embodiments, the NLS is fused to the N-terminus of the adenosine deaminase. In some embodiments, the NLS is fused to the C-terminus of the adenosine deaminase. In some embodiments, the NLS is fused to the fusion protein via one or more linkers. In some embodiments, the NLS is fused to the fusion protein without a linker. In some embodiments, the NLS comprises an amino acid sequence of any one of the NLS sequences provided or referenced herein. In some embodiments, the NLS comprises an amino acid sequence as set forth in SEQ ID NO: 4 or SEQ ID NO: 5. Additional 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.*, PCT/EP2000/011690, the contents of which are incorporated herein by reference for their disclosure of exemplary nuclear localization sequences. In some embodiments, a NLS comprises the amino acid sequence PKKKRKV (SEQ ID NO: 4) or MDSLLMNRRKFLYQFKNVRWAKGRRETYLC (SEQ ID NO: 5).

[00325] In some embodiments, the general architecture of exemplary fusion proteins with an adenosine deaminase and a napDNAbp comprises any one of the following structures, where NLS is a nuclear localization sequence (*e.g.,* any NLS provided herein), NH_2 is the N-terminus of the fusion protein, and COOH is the C-terminus of the fusion protein.

[00326] Fusion proteins comprising an adenosine deaminase, a napDNAbp, and a NLS.

NH_2-[NLS]-[adenosine deaminase]-[napDNAbp]-COOH;

NH_2-[adenosine deaminase]-[NLS]-[napDNAbp]-COOH;

NH_2-[adenosine deaminase]-[napDNAbp]-[NLS]-COOH;

NH_2-[NLS]-[napDNAbp]-[adenosine deaminase]-COOH;

NH_2-[napDNAbp]-[NLS]-[adenosine deaminase]-COOH;

NH_2-[napDNAbp]-[adenosine deaminase]-[NLS]-COOH;

[00327] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between one or more of the domains or proteins (*e.g.,* adenosine deaminase, napDNAbp, and/or NLS). In some embodiments, the "-

" used in the general architecture above indicates the presence of an optional linker.

[00328] Fusion proteins comprising an adenosine deaminase, a napDNAbp, and an inhibitor of base repair (IBR).

NH_2-[IBR]-[adenosine deaminase]-[napDNAbp]-COOH;

NH_2-[adenosine deaminase]-[IBR]-[napDNAbp]-COOH;

NH_2-[adenosine deaminase]-[napDNAbp]-[IBR]-COOH;

NH_2-[IBR]-[napDNAbp]-[adenosine deaminase]-COOH;

NH_2-[napDNAbp]-[IBR]-[adenosine deaminase]-COOH;

NH_2-[napDNAbp]-[adenosine deaminase]-[IBR]-COOH;

[00329] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between one or more of the domains or proteins (*e.g.*, adenosine deaminase, napDNAbp, and/or IBR). In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker.

[00330] Fusion proteins comprising an adenosine deaminase, a napDNAbp,an inhibitor of base repair (IBR) and a NLS.

NH_2-[IBR]-[NLS]-[adenosine deaminase]-[napDNAbp]-COOH;

NH_2-[NLS]-[IBR]-[adenosine deaminase]-[napDNAbp]-COOH;

NH_2-[NLS]-[adenosine deaminase]-[IBR]-[napDNAbp]-COOH;

NH_2-[NLS]-[adenosine deaminase]-[napDNAbp]-[IBR]-COOH;

NH_2-[IBR]-[adenosine deaminase]-[NLS]-[napDNAbp]-COOH;

NH_2-[adenosine deaminase]-[IBR]-[NLS]-[napDNAbp]-COOH;

NH_2-[adenosine deaminase]-[NLS]-[IBR]-[napDNAbp]-COOH;

NH_2-[adenosine deaminase]-[NLS]-[napDNAbp]-[IBR]-COOH;

NH_2-[IBR]-[adenosine deaminase]-[napDNAbp]-[NLS]-COOH;

NH_2-[adenosine deaminase]-[IBR]-[napDNAbp]-[NLS]-COOH;

NH_2-[adenosine deaminase]-[napDNAbp]-[IBR]-[NLS]-COOH;

NH_2-[adenosine deaminase]-[napDNAbp]-[NLS]-[IBR]-COOH;

NH_2-[IBR]-[NLS]-[napDNAbp]-[adenosine deaminase]-COOH;

NH_2-[NLS]-[IBR]-[napDNAbp]-[adenosine deaminase]-COOH;

NH_2-[NLS]-[napDNAbp]-[IBR]-[adenosine deaminase]-COOH;

NH_2-[NLS]-[napDNAbp]-[adenosine deaminase]-[IBR]-COOH;

NH_2-[IBR]-[napDNAbp]-[NLS]-[adenosine deaminase]-COOH;

NH_2-[napDNAbp]-[IBR]-[NLS]-[adenosine deaminase]-COOH;

NH_2-[napDNAbp]-[NLS]-[IBR]-[adenosine deaminase]-COOH;

NH_2-[napDNAbp]-[NLS]-[adenosine deaminase]-[IBR]-COOH;

NH_2-[IBR]-[napDNAbp]-[adenosine deaminase]-[NLS]-COOH;

NH_2-[napDNAbp]-[IBR]-[adenosine deaminase]-[NLS]-COOH;

NH_2-[napDNAbp]-[adenosine deaminase]-[IBR]-[NLS]-COOH;

NH_2-[napDNAbp]-[adenosine deaminase]-[NLS]-[IBR]-COOH;

[00331] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between one or more of the domains or proteins (*e.g.,* adenosine deaminase, napDNAbp, NLS, and/or IBR). In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker.

[00332] Some aspects of the disclosure provide fusion proteins that comprise a nucleic acid programmable DNA binding protein (napDNAbp) and at least two adenosine deaminase domains. Without wishing to be bound by any particular theory, dimerization of adenosine deaminases (*e.g.,* in cis or in trans) may improve the ability (*e.g.,* efficiency) of the fusion protein to modify a nucleic acid base, for example to deaminate adenine. In some embodiments, any of the fusion proteins may comprise 2, 3, 4 or 5 adenosine deaminase domains. In some embodiments, any of the fusion proteins provided herein comprise two adenosine deaminases. In some embodiments, any of the fusion proteins provided herein contain only two adenosine deaminases. In some embodiments, the adenosine deaminases are the same. In some embodiments, the adenosine deaminases are any of the adenosine deaminases provided herein. In some embodiments, the adenosine deaminases are different. In some embodiments, the first adenosine deaminase is any of the adenosine deaminases provided herein, and the second adenosine is any of the adenosine deaminases provided herein, but is not identical to the first adenosine deaminase. As one example, the fusion protein may comprise a first adenosine deaminase and a second adenosine deaminase that both comprise the amino acid sequence of SEQ ID NO: 72, which contains a A106V, D108N, D147Y, and E155V mutation from ecTadA (SEQ ID NO: 1). As another example, the fusion protein may comprise a first adenosine deaminase domain that comprises the amino amino acid sequence of SEQ ID NO: 72, which contains a A106V, D108N, D147Y, and E155V mutation from ecTadA (SEQ ID NO: 1), and a second adenosine deaminase that comprises the amino acid sequence of SEQ ID NO: 421, which contains a L84F, A106V, D108N, H123Y, D147Y, E155V, and I156F mutation from ecTadA (SEQ ID NO: 1).

[00333] In some embodiments, the fusion protein comprises two adenosine deaminases (*e.g.,* a first adenosine deaminase and a second adenosine deaminase). In some embodiments, the fusion protein comprises a first adenosine deaminase and a second adenosine deaminase.

In some embodiments, the first adenosine deaminase is N-terminal to the second adenosine deaminase in the fusion protein. In some embodiments, the first adenosine deaminase is C-terminal to the second adenosine deaminase in the fusion protein. In some embodiments, the first adenosine deaminase and the second deaminase are fused directly or via a linker. In some embodiments, the linker is any of the linkers provided herein, for example, any of the linkers described in the "Linkers" section. In some embodiments, the linker comprises the amino acid sequence of any one of SEQ ID NOs: 10, 37-40, 384-386, or 685-688. In some embodiments, the first adenosine deaminase is the same as the second adenosine deaminase. In some embodiments, the first adenosine deaminase and the second adenosine deaminase are any of the adenosine deaminases described herein. In some embodiments, the first adenosine deaminase and the second adenosine deaminase are different. In some embodiments, the first adenosine deaminase is any of the adenosine deaminases provided herein. In some embodiments, the second adenosine deaminase is any of the adenosine deaminases provided herein but is not identical to the first adenosine deaminase. In some embodiments, the first adenosine deaminase is an ecTadA adenosine deaminase. In some embodiments, the first adenosine deaminase comprises an amino acid sequence that is at least 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% identical to any one of the amino acid sequences set forth in any one of SEQ ID NOs: 1, 64-84, 420-437, 672-684, or to any of the adenosine deaminases provided herein. In some embodiments, the first adenosine deaminase comprises the amino acid sequence of SEQ ID NO: 1. In some embodiments, the second adenosine deaminase comprises an amino acid sequence that is at least 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% identical to any one of the amino acid sequences set forth in any one of SEQ ID NOs: 1, 64-84, 420-437, 672-684, or to any of the adenosine deaminases provided herein. In some embodiments, the second adenosine deaminase comprises the amino acid sequence of SEQ ID NO: 1. In some embodiments, the first adenosine deaminase and the second adenosine deaminase of the fusion protein comprise the mutations in ecTadA (SEQ ID NO: 1), or corresponding mutations in another adenosine deaminase, as shown in any one of the constructs provided in Table 4 (*e.g.*, pNMG-371, pNMG-477, pNMG-576, pNMG-586, and pNMG-616). In some embodiments, the fusion protein comprises the two adenosine deaminases (*e.g.*, a first adenosine deaminase and a second adenosine deaminase) of any one of the constructs (*e.g.*, pNMG-371, pNMG-477, pNMG-576, pNMG-586, and pNMG-616) in Table 4.

[00334] In some embodiments, the general architecture of exemplary fusion proteins with a first adenosine deaminase, a second adenosine deaminase, and a napDNAbp comprises any one of the following structures, where NLS is a nuclear localization sequence (*e.g.,* any NLS provided herein), NH_2 is the N-terminus of the fusion protein, and COOH is the C-terminus of the fusion protein.

[00335] Fusion proteins comprising a first adenosine deaminase, a second adenosine deaminase, and a napDNAbp.

NH_2-[first adenosine deaminase]-[second adenosine deaminase]-[napDNAbp]-COOH;
NH_2-[first adenosine deaminase]-[napDNAbp]-[second adenosine deaminase]-COOH;
NH_2-[napDNAbp]-[first adenosine deaminase]-[second adenosine deaminase]-COOH;
NH_2-[second adenosine deaminase]-[first adenosine deaminase]-[napDNAbp]-COOH;
NH_2-[second adenosine deaminase]-[napDNAbp]-[first adenosine deaminase]-COOH;
NH_2-[napDNAbp]-[second adenosine deaminase]-[first adenosine deaminase]-COOH;

[00336] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between one or more of the domains or proteins (*e.g.,* first adenosine deaminase, second adenosine deaminase, and/or napDNAbp). In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker.

[00337] Fusion proteins comprising a first adenosine deaminase, a second adenosine deaminase, a napDNAbp, and an NLS.

NH_2-[NLS]-[first adenosine deaminase]-[second adenosine deaminase]-[napDNAbp]-COOH;
NH_2-[first adenosine deaminase]-[NLS]-[second adenosine deaminase]-[napDNAbp]-COOH;
NH_2-[first adenosine deaminase]-[second adenosine deaminase]-[NLS]-[napDNAbp]-COOH;
NH_2-[first adenosine deaminase]-[second adenosine deaminase]-[napDNAbp]-[NLS]-COOH;
NH_2-[NLS]-[first adenosine deaminase]-[napDNAbp]-[second adenosine deaminase]-COOH;
NH_2-[first adenosine deaminase]-[NLS]-[napDNAbp]-[second adenosine deaminase]-COOH;
NH_2-[first adenosine deaminase]-[napDNAbp]-[NLS]-[second adenosine deaminase]-COOH;
NH_2-[first adenosine deaminase]-[napDNAbp]-[second adenosine deaminase]-[NLS]-COOH;
NH_2-[NLS]-[napDNAbp]-[first adenosine deaminase]-[second adenosine deaminase]-COOH;
NH_2-[napDNAbp]-[NLS]-[first adenosine deaminase]-[second adenosine deaminase]-COOH;
NH_2-[napDNAbp]-[first adenosine deaminase]-[NLS]-[second adenosine deaminase]-COOH;
NH_2-[napDNAbp]-[first adenosine deaminase]-[second adenosine deaminase]-[NLS]-COOH;
NH_2-[NLS]-[second adenosine deaminase]-[first adenosine deaminase]-[napDNAbp]-COOH;
NH_2-[second adenosine deaminase]-[NLS]-[first adenosine deaminase]-[napDNAbp]-COOH;

NH_2-[second adenosine deaminase]-[first adenosine deaminase]-[NLS]-[napDNAbp]-COOH;
NH_2-[second adenosine deaminase]-[first adenosine deaminase]-[napDNAbp]-[NLS]-COOH;
NH_2-[NLS]-[second adenosine deaminase]-[napDNAbp]-[first adenosine deaminase]-COOH;
NH_2-[second adenosine deaminase]-[NLS]-[napDNAbp]-[first adenosine deaminase]-COOH;
NH_2-[second adenosine deaminase]-[napDNAbp]-[NLS]-[first adenosine deaminase]-COOH;
NH_2-[second adenosine deaminase]-[napDNAbp]-[first adenosine deaminase]-[NLS]-COOH;
NH_2-[NLS]-[napDNAbp]-[second adenosine deaminase]-[first adenosine deaminase]-COOH;
NH_2-[napDNAbp]-[NLS]-[second adenosine deaminase]-[first adenosine deaminase]-COOH;
NH_2-[napDNAbp]-[second adenosine deaminase]-[NLS]-[first adenosine deaminase]-COOH;
NH_2-[napDNAbp]-[second adenosine deaminase]-[first adenosine deaminase]-[NLS]-COOH;

[00338] In some embodiments, the fusion proteins provided herein do not comprise a linker. In some embodiments, a linker is present between one or more of the domains or proteins (*e.g.,* first adenosine deaminase, second adenosine deaminase, napDNAbp, and/or NLS). In some embodiments, the "-" used in the general architecture above indicates the presence of an optional linker.

[00339] It should be appreciated that the fusion proteins of the present disclosure may comprise one or more additional features. For example, in some embodiments, the fusion protein may comprise 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. 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.

Linkers

[00340] In certain embodiments, linkers may be used to link any of the protein or protein domains described herein. 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.

[00341] 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 a bond (*e.g.*, a covalent bond), 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, 100-110, 110-120, 120-130, 130-140, 140-150, or 150-200 amino acids in length. Longer or shorter linkers are also contemplated. In some embodiments, a linker comprises the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 10), which may also be referred to as the XTEN linker. In some embodiments, a linker comprises the amino acid sequence SGGS (SEQ ID NO: 37). In some embodiments, a linker comprises $(SGGS)_n$ (SEQ ID NO: 37), (GGGS)n (SEQ ID NO: 38), (GGGGS)n (SEQ ID NO: 39), (G)n, (EAAAK)n (SEQ ID NO: 40), (GGS)n, SGSETPGTSESATPES (SEQ ID NO: 10), or $(XP)_n$ motif, or a combination of any of these, wherein n is independently an integer between 1 and 30, and wherein X is any amino acid. In some embodiments, n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15. In some embodiments, a linker comprises SGSETPGTSESATPES (SEQ ID NO: 10), and SGGS (SEQ ID NO: 37). In some embodiments, a linker comprises

SGGSSGSETPGTSESATPESSGGS (SEQ ID NO: 384). In some embodiments, a linker comprises SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 385). In some embodiments, a linker comprises GGSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSGGS (SEQ ID NO: 386). In some embodiments, the linker is 24 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPES (SEQ ID NO: 685). In some embodiments, the linker is 40 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS (SEQ ID NO: 686). In some embodiments, the linker is 64 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 687). In some embodiments, the linker is 92 amino acids in length. In some embodiments, the linker comprises the amino acid sequence PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATS (SEQ ID NO: 688). It should be appreciated that any of the linkers provided herein may be used to link a first adenosine deaminase and a second adenosine deaminase; an adenosine deaminase (*e.g.*, a first or a second adenosine deaminase) and a napDNAbp; a napDNAbp and an NLS; or an adenosine deaminase (*e.g.*, a first or a second adenosine deaminase) and an NLS.

[00342] In some embodiments, any of the fusion proteins provided herein, comprise an adenosine deaminase and a napDNAbp that are fused to each other via a linker. In some embodiments, any of the fusion proteins provided herein, comprise a first adenosine deaminase and a second adenosine deaminase that are fused to each other via a linker. In some embodiments, any of the fusion proteins provided herein, comprise an NLS, which may be fused to an adenosine deaminase (*e.g.*, a first and/or a second adenosine deaminase), a nucleic acid programmable DNA binding protein (napDNAbp), and or an inhibitor of base repair (IBR). Various linker lengths and flexibilities between an adenosine deaminase (*e.g.*, an engineered ecTadA) and a napDNAbp (*e.g.,* a Cas9 domain), and/or between a first adenosine deaminase and a second adenosine deaminase can be employed (*e.g.,* ranging from very flexible linkers of the form (GGGGS)n (SEQ ID NO: 38), (GGGGS)n (SEQ ID NO: 39), and (G)n to more rigid linkers of the form (EAAAK)n (SEQ ID NO: 40), (SGGS)n (SEQ ID NO: 37), SGSETPGTSESATPES (SEQ ID NO: 10) (see, *e.g.*, Guilinger JP, Thompson

DB, Liu DR. Fusion of catalytically inactive Cas9 to FokI nuclease improves the specificity of genome modification. *Nat. Biotechnol.* 2014; 32(6): 577-82; the entire contents are incorporated herein by reference) and $(XP)_n$) in order to achieve the optimal length for deaminase activity for the specific application. In some embodiments, n is 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, or 15. In some embodiments, the linker comprises a $(GGS)_n$ motif, wherein n is 1, 3, or 7. In some embodiments, the adenosine deaminase and the napDNAbp, and/or the first adenosine deaminase and the second adenosine deaminase of any of the fusion proteins provided herein are fused via a linker comprising the amino acid sequence SGSETPGTSESATPES (SEQ ID NO: 10), SGGS (SEQ ID NO: 37), SGGSSGSETPGTSESATPESSGGS (SEQ ID NO: 384), SGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 385), or GGSGGSPGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATSGGSGGS (SEQ ID NO: 386). In some embodiments, the linker is 24 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPES (SEQ ID NO: 685). In some embodiments, the linker is 40 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGS (SEQ ID NO: 686). In some embodiments, the linker is 64 amino acids in length. In some embodiments, the linker comprises the amino acid sequence SGGSSGGSSGSETPGTSESATPESSGGSSGGSSGGSSGGSSGSETPGTSESATPESSGGSSGGS (SEQ ID NO: 687). In some embodiments, the linker is 92 amino acids in length. In some embodiments, the linker comprises the amino acid sequence PGSPAGSPTSTEEGTSESATPESGPGTSTEPSEGSAPGSPAGSPTSTEEGTSTEPSEGSAPGTSTEPSEGSAPGTSESATPESGPGSEPATS (SEQ ID NO: 688).

[00343] Some aspects of the disclosure provide fusion proteins comprising a Cas9 domain and an adenosine deaminase. Exemplary fusion proteins include, without limitation, the following fusion proteins (for the purposes of clarity, the adenosine deaminase domain is shown in **Bold**; mutations of the ecTadA deaminase domain are shown in **Bold underlining**; the XTEN linker is shown in *italics*; the UGI/AAG/EndoV domains are shown in ***Bold italics***; and NLS is shown in *underlined italics*):

ecTadA(wt)-XTEN-nCas9-NLS:

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH

DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 11)

ecTadA(D108N)-XTEN-nCas9-NLS: (mammalian construct, active on DNA, A to G editing):

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV

LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI
EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR
IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY
GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE
TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL
EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 12)

ecTadA(D108G)-XTEN-nCas9-NLS: (mammalian construct, active on DNA, A to G editing):

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARGAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA

KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI
EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR
IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY
GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE
TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL
EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 13)

ecTadA(D108V)-XTEN-nCas9-NLS: (mammalian construct, active on DNA, A to G editing):

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARVAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK

MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI
EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR
IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY
GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE
TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL
EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 14)

ecTadA(D108N)-XTEN-nCas9-UGI-NLS (BE3 analog of A to G editor):
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI
EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN

EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR
IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY
GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE
TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL
EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGDSGGS***TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHT***
AYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKMLSGGS*PKKKRKV* (SEQ ID
NO: 15)

ecTadA(D108G)-XTEN-nCas9-UGI-NLS (BE3 analog of A to G editor):
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARGAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI
EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL

AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS*TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHT AYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML*SGGS*PKKKRKV* (SEQ ID NO: 16)

ecTadA(D108V)-XTEN-nCas9-UGI-NLS (BE3 analog of A to G editor):
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARVAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH**

IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS*TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHT AYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML*SGGS*PKKKRKV* (SEQ ID NO: 17)

ecTadA(D108N)-XTEN-dCas9-UGI-NLS (mammalian cells, BE2 analog of A to G editor):
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDA IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV**

KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS*TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHT AYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML*SGGS*PKKKRKV* (SEQ ID NO: 18)

ecTadA(D108G)-XTEN-dCas9-UGI-NLS (mammalian cells, BE2 analog of A to G editor):
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARGAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDA IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE

TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS***TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHT AYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML***SGGS*PKKKRKV* (SEQ ID NO: 19)

ecTadA(D108V)-XTEN-dCas9-UGI-NLS (mammalian cells, BE2 analog of A to G editor):
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARVAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDA IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL

EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS***TNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESDILVHT AYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML***SGGS*PKKKRKV* (SEQ ID NO: 20)

ecTadA(D108N)-XTEN-nCas9-AAG(E125Q)-NLS – cat. alkyladenosine glycosylase:
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH

RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS***KGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELRGRI VETQAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDGA CVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQR DLAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVD RVAEQDTQA***SGGS***PKKKRKV*** (SEQ ID NO: 21)

ecTadA(D108G)-XTEN-nCas9-AAG(E125Q)-NLS – cat. alkyladenosine glycosylase:
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARGAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH

RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS***KGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELRGRI VETQAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDGA CVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQR DLAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVD RVAEQDTQA***SGGS***PKKKRKV*** (SEQ ID NO: 22)

ecTadA(D108V)-XTEN-nCas9-AAG(E125Q)-NLS – cat. alkyladenosine glycosylase:
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARVAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH

RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS***KGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELRGRI VETQAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDGA CVLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQR DLAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVD RVAEQDTQA***SGGS***PKKKRKV*** (SEQ ID NO: 23)

ecTadA(D108N)-XTEN-nCas9-EndoV(D35A)-NLS: contains cat. endonuclease V :

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH

RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS***DLASLRAQQIELASSVIREDRLDKDPPDLIAGAAVGFEQGGEV TRAAMVLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVF VDGHGISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQ LAWVWRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPA FVRYTANQP***SGGS*PKKKRKV* (SEQ ID NO: 24)

ecTadA(D108G)-XTEN-nCas9-EndoV (D35A)-NLS: contains cat. endonuclease V:
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARGAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH

RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS***DLASLRAQQIELASSVIREDRLDKDPPDLIAGAAVGFEQGGEV TRAAMVLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVF VDGHGISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQ LAWVWRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPA FVRYTANQP***SGGS*PKKKRKV* (SEQ ID NO: 25)

ecTadA(D108V)-XTEN-nCas9-EndoV(D35A)-NLS: contains cat. endonuclease V:
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARVAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH

RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDSGGS***DLASLRAQQIELASSVIREDRLDKDPPDLIAGAAVGFEQGGEV TRAAMVLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVF VDGHGISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQ LAWVWRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPA FVRYTANQP***SGGS*PKKKRKV* (SEQ ID NO: 26)

Variant resulting from first round of evolution (in bacteria) ecTadA(H8Y_D108N_N127S)-XTEN-dCas9:

MSEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARNAKTGAAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDA IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA

SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGD (SEQ ID NO: 27)

Enriched variants from second round of evolution (in bacteria) ecTadA
(H8Y_D108N_N127S_E155X)-XTEN-dCas9; X=D, G or V:

**MSEVEFSYEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARNAKTGAAGSLMDVLHHPGMSHRVEITEGILADECAALLSDFFRMRRQXIKA
QKKAQSSTD***SGSETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV
LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK
VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD
LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA
KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS
KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK
MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI
EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF
DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK
TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN
EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI
RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL
AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR
IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDA
IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK
FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV
KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY
GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE
TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD
WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL
EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA
SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH
RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE
TRIDLSQLGGD (SEQ ID NO: 28)

ecTadA*-XTEN-nCas9-GGS-DNA repair inhibitor-GGS-NLS (Inhibitor = UGI, AAG*E125Q or EndoV*D35A)

pNMG-160: **ecTadA(D108N)**-*XTEN*-nCas9-GGS-***AAG*(E125Q)***-GGS-*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE TRIDLSQLGGDGGS***KGHLTRLGLEFFDQPAVPLARAFLGQVLVRRLPNGTELRGRIV***

ETQAYLGPEDEAAHSRGGRQTPRNRGMFMKPGTLYVYIIYGMYFCMNISSQGDGAC VLLRALEPLEGLETMRQLRSTLRKGTASRVLKDRELCSGPSKLCQALAINKSFDQRD LAQDEAVWLERGPLEPSEPAVVAAARVGVGHAGEWARKPLRFYVRGSPWVSVVDR VAEQDTQAGGS*PKKKRKV* (SEQ ID NO: 387)

pNMG-161: **ecTadA(D108N)**-*XTEN*-nCas9-GGS-***EndoV*(D35A)***-GGS-*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGSETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKV LGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAK VDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKAD LRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDA KAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLS KDTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEK MDGTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKI EKILTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNF DKNLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFK TNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEEN EDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGI RDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANL AGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKR IEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDH IVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRK FDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREV KVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVY GDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGE TGEIVWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKD WDPKKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFL EAKGYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLA SHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKH RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYE*

TRIDLSQLGGDGGS***DLASLRAQQIELASSVIREDRLDKDPPDLIAGAAVGFEQGGEVT RAAMVLLKYPSLELVEYKVARIATTMPYIPGFLSFREYPALLAAWEMLSQKPDLVFV DGHGISHPRRLGVASHFGLLVDVPTIGVAKKRLCGKFEPLSSEPGALAPLMDKGEQL AWVWRSKARCNPLFIATGHRVSVDSALAWVQRCMKGYRLPEPTRWADAVASERPA FVRYTANQP***GGS*PKKKRKV* (SEQ ID NO: 388)

pNMG-371: **ecTadA(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)**-SGGS-SGGS-*XTEN*-SGGS-SGGS-**ecTadA(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)**-SGGS-SGGS-*XTEN*-SGGS-SGGS-nCas9-SGGS-*NLS*

SEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHD PTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGV RNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQ KKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALTL AKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLVM QNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHY PGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTD**SGGSSGGS*SGS ETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNT DRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDS FFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLI YLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAIL SARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDT YDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEH HQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDG TEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKIL TFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKN LPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRK VTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILE DIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQ SGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIK ELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSF**

LKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTK
AERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS
KLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVY
DVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWD
KGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKY
GGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKE
VKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLK
GSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIRE
QAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQ
LGGDSGGS*PKKKRKV* (SEQ ID NO: 440)

pNMG-616 amino acid sequence: **ecTadA**$_{\text{(wild type)}}$-(SGGS)2-*XTEN*-(SGGS)2-
ecTadA$_{\text{(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F _K157N)}}$-
(SGGS)2-*XTEN*-(SGGS)2_nCas9_SGGS_NLS
MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT**
LAKRALDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV
MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGS*S*
*GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD
SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI
LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI
LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG

SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 691)

pNMG-624 amino acid sequence: **ecTadA**$_{\text{(wild type)}}$-*32 a.a. linker-*
ecTadA$_{\text{(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)}}$-*24 a.a. linker*_nCas9_SGGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGGSSGGSSGSETPGTSESATPESSGGSSGGS***SEVEFSHEYWMRHALTL AKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVM QNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHY PGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD***SGGSSGGSSGS ETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNL IGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI KFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLL KALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL

FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF
LKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT
VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILK
EHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSID
NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL
SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDF
RKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM
IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI
IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH
LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSG
GS*PKKKRK*V (SEQ ID NO: 692)

pNMG-476 amino acid sequence (evolution #3 hetero dimer, wt TadA + TadA evo #3 mutations): **ecTadA$_{(wild\text{-}type)}$**-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA$_{(L84F_A106V_D108N_H123Y_D147Y_E155V_I156F)}$**-(SGGS)2-*XTEN*-(SGGS)2_nCas9_SGGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT**
LAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRHDPTAHAEIMALRQGGLV
MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNHRVEITEGILADECAALLSYFFRMRRQVFKAQKKAQSSTDSGGSSGGS*S*
*GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD
SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI

LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI
LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK
GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI
DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 693)

pNMG-477 amino acid sequence: **ecTadA$_{(wild-type)}$**-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA$_{(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)}$**-(SGGS)2-*XTEN*-(SGGS)2_nCas9_SGGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT**
LAKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGLHDPTAHAEIMALRQGGLV
MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTDSGGSSGGS*S*
*GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD
SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE

HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI
LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI
LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK
GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI
DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 694)

pNMG-558 amino acid sequence: **ecTadA$_{(wild-type)}$**- *32 a.a. linker-*
ecTadA$_{(H36L_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)}$- *24 a.a.*
linker_nCas9_SGGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTD*SGGSSGGSSGSETPGTSESATPESSGGSSGGS***SEVEFSHEYWMRHALTL**
AKRAWDEREVPVGAVLVLNNRVIGEGWNRPIGLHDPTAHAEIMALRQGGLVM
QNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHY
PGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTD*SGGSSGGSSGS*
*ETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNL
IGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF
LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI
KFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR

RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN
LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLL
KALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL
NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV
GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP
KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL
FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF
LKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT
VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILK
EHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSID
NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL
SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDF
RKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM
IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI
IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH
LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSG
GS*PKKKRKV* (SEQ ID NO: 695)

pNMG-576 amino acid sequence: **ecTadA$_{(wild-type)}$**-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA$_{(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)}$**-(SGGS)2-*XTEN*-(SGGS)2_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT**
LAKRAWDEREVPVGAVLVLNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLV
MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTDSGGSSGGS*S*
*GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD

SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI
LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI
LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK
GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI
DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 696)

pNMG-577 amino acid sequence: **ecTadA$_{(wild-type)}$**-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA$_{(H36L_P48S_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F_K157N)}$**-(SGGS)2-*XTEN*-(SGGS)2_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT**
LAKRAWDEREVPVGAVLVLNNRVIGEGWNRSIGLHDPTAHAEIMALRQGGLV
MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNHRVEITEGILADECNALLCYFFRMRRQVFNAQKKAQSSTDSGGSSGGS*S*

*GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD
SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI
LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI
LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK
GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI
DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 697)

pNMG-586 amino acid sequence: **ecTadA$_{(wild-type)}$**-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA$_{(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_E155V_I156F_K157N)}$**-(SGGS)2-*XTEN*-(SGGS)2_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT**
LAKRAWDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV

MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH YPGMNHRVEITEGILADECAALLCYFFRMRRQVFNAQKKAQSSTDSGGSSGGS*S GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 698)

pNMG-588 amino acid sequence: **ecTadA**$_{\text{(wild-type)}}$-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA**$_{\text{(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_A142N_D147Y_E155V_I156F_K157N)}}$-(SGGS)2-*XTEN*-(SGGS)2_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA

QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT LAKRAWDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH YPGMNHRVEITEGILADECNALLCYFFRMRRQVFNAQKKAQSSTD**SGGSSGGS*S GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 699)

pNMG-620 amino acid sequence: **ecTadA**$_{\text{(wild-type)}}$-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA**$_{\text{(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)}}$-(SGGS)2-*XTEN*-(SGGS)2_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH

DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT LAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH YPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD**SGGSSGGS*S GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 700)

pNMG-617 amino acid sequence: **ecTadA*(wild-type)***-(SGGS)2-*XTEN*-(SGGS)2-

ecTadA$_{(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_E155V_I156F_K157N)^-}$

(SGGS)2-*XTEN*-(SGGS)2_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT**
LAKRALDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV
MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNHRVEITEGILADECNALLCYFFRMRRQVFNAQKKAQSSTDSGGSSGGS*S*
*GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD
SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI
LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI
LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK
GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI
DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 701)

pNMG-618 amino acid sequence: **ecTadA$_{(wild-type)}$**-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA$_{(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142A_S146C_D147Y_R152P_E155V_I156F_K157N)}$**-(SGGS)2-*XTEN*-(SGGS)2_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT**
LAKRALDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV
MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH
YPGMNHRVEITEGILADECNALLCYFFRMPRQVFNAQKKAQSSTDSGGSSGGS*S*
*GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN
TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD
SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL
IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI
LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD
TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE
HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD
GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI
LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK
NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR
KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI
LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD
KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG
SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE
GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP
QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN
LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT
LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY
KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI
VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP
KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK
GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY
EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK
PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI

DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 702)

pNMG-620 amino acid sequence: **ecTadA$_{(wild-type)}$**-(SGGS)2-*XTEN*-(SGGS)2-**ecTadA$_{(W23R_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)}$**-(SGGS)2-*XTEN*-(SGGS)2_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTDSGGSSGGS*SGSETPGTSESATPES*SGGSSGGS**SEVEFSHEYWMRHALT LAKRARDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLV MQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLH YPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD**SGGSSGGS*S GSETPGTSESATPES*SGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGN TDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDD SFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRL IYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAI LSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDE HHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMD GTEELLVKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKI LTFRIPYYVGPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDK NLPNEKVLPKHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNR KVTVKQLKEDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDI LEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRD KQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAG SPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEE GIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVP QSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDN LTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVIT LKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDY KVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEI VWDKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDP KKYGGFDSPTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAK GYKEVKKDLIIKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHY

EKLKGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDK PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRI DLSQLGGDSGGS*PKKKRKV* (SEQ ID NO: 703)

pNMG-621 amino acid sequence: **ecTadA$_{(wild-type)}$**- *32 a.a. linker-* **ecTadA$_{(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)}$**- *24 a.a. linker*_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGGSSGGSSGSETPGTSESATPESSGGSSGGS***SEVEFSHEYWMRHALTL AKRAWDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVM QNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHY PGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD***SGGSSGGSSGS ETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNL IGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI KFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLL KALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF LKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILK EHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSID NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDF RKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP

TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSG GS*PKKKRKV* (SEQ ID NO: 704)

pNMG-622 amino acid sequence: **ecTadA**$_{(wild\text{-}type)}$- *32 a.a. linker-* **ecTadA**$_{(H36L_P48A_R51L_L84F_A106V_D108N_H123Y_A142N_S146C_D147Y_R152P_E155V_I156F_K157N)}$- *24 a.a. linker*_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA QKKAQSSTD*SGGSSGGSSGSETPGTSESATPESSGGSSGGS***SEVEFSHEYWMRHALTL AKRAWDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVM QNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHY PGMNHRVEITEGILADECNALLCYFFRMPRQVFNAQKKAQSSTD***SGGSSGGSSGS ETPGTSESATPES***DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNL IGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI KFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLL KALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF LKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILK EHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSID NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDF RKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM

IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF
ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP
TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI
IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN
EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH
LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSG
GS*PKKKRKV* (SEQ ID NO: 705)

pNMG-623 amino acid sequence: **ecTadA**$_{(wild\text{-}type)}$- *32 a.a. linker-*
ecTadA$_{(W23L_H36L_P48A_R51L_L84F_A106V_D108N_H123Y_S146C_D147Y_R152P_E155V_I156F_K157N)}$- *24 a.a. linker*_nCas9_GGS_*NLS*

MSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEGWNRPIGRH
DPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFG
ARDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKA
QKKAQSSTD*SGGSSGGSSGSETPGTSESATPESSGGSSGGS***SEVEFSHEYWMRHALTL**
AKRALDEREVPVGAVLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVM
QNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGVRNAKTGAAGSLMDVLHY
PGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTD*SGGSSGGSSGS*
*ETPGTSESATPES*DKKYSIGLAIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNL
IGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVDDSFFHRLEESF
LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYLALAHMI
KFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSR
RLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDN
LLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLL
KALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKL
NREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYV
GPLARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLP
KHSLLYEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLK
EDYFKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTL
FEDREMIEERLKTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDF
LKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQT
VKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILK
EHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSID
NKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNLTKAERGGL

SELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKLVSDF RKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKM IAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDF ATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSP TVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLI IKLPKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDN EQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIH LFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSG GS*PKKKRKV* (SEQ ID NO: 706)

[00344] In some embodiments, the fusion protein comprises an amino acid sequence that is 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% identical to any one of the amino acid sequences set forth in any one of SEQ ID NOs: 11-28, 387, 388, 440, 691-706, or to any of the fusion proteins provided herein. In some embodiments, the fusion protein comprises an amino acid sequence that has 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 mutations compared to any one of the amino acid sequences set forth in SEQ ID NOs: 11-28, 387, 388, 440, 691-706, or any of the fusion proteins provided herein. In some embodiments, the fusion protein comprises an amino acid sequence that has at least 5, at least 10, at least 15, at least 20, at least 25, at least 30, at least 35, at least 40, at least 45, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 110, at least 120, at least 130, at least 140, at least 150, at least 160, at least 170, at least 200, at least 300, at least 400, at least 500, at least 600, at least 700, at least 800, at least 900, at least 1000, at least 1100, at least 1200, at least 1300, at least 1400, at least 1500, at least 1600, at least 1700, at least 1750, or at least 1800 identical contiguous amino acid residues as compared to any one of the amino acid sequences set forth in SEQ ID NOs: 11-28, 387, 388, 440, 691-706, or any of the fusion proteins provided herein.

Nucleic acid programmable DNA binding protein (napDNAbp) complexes with guide nucleic acids

[00345] Some aspects of this disclosure provide complexes comprising any of the fusion proteins provided herein, and a guide nucleic acid bound to napDNAbp of the fusion

protein. Some aspects of this disclosure provide complexes comprising any of the fusion proteins provided herein, and a guide RNA bound to a Cas9 domain (*e.g.*, a dCas9, a nuclease active Cas9, or a Cas9 nickase) of fusion protein.

[00346] In some embodiments, the guide nucleic acid (*e.g.*, guide RNA) is from 15-100 nucleotides long and comprises a sequence of at least 10 contiguous nucleotides that is complementary to a target sequence. In some embodiments, the guide RNA is 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 nucleotides long. In some embodiments, the guide RNA comprises a sequence of 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 that is complementary to a target sequence. In some embodiments, the target sequence is a DNA sequence. In some embodiments, the target sequence is an RNA sequence. In some embodiments, the target sequence is a sequence in the genome of a mammal. In some embodiments, the target sequence is a sequence in the genome of a human. In some embodiments, the 3′ end of the target sequence is immediately adjacent to a canonical PAM sequence (NGG). In some embodiments, the guide nucleic acid (*e.g.*, guide RNA) is complementary to a sequence associated with a disease or disorder. In some embodiments, the guide nucleic acid (*e.g.*, guide RNA) is complementary to a sequence associated with a disease or disorder having a mutation in a gene selected from the genes disclosed in any one of Tables 1 and 2.

Methods of using fusion proteins comprising an adenosine deaminase and a nucleic acid programmable DNA binding protein (napDNAbp) domain

[00347] Some aspects of this disclosure provide methods of using the fusion proteins, or complexes comprising a guide nucleic acid (*e.g.*, gRNA) and a nucleobase editor provided herein. For example, some aspects of this disclosure provide methods comprising contacting a DNA, or RNA molecule with any of the fusion proteins provided herein, and with at least one guide nucleic acid (*e.g.*, guide RNA), wherein the guide nucleic acid, (*e.g.*, guide RNA) is about 15-100 nucleotides long and comprises a sequence of at least 10 contiguous nucleotides that is complementary to a target sequence. In some embodiments, the 3' end of the target sequence is immediately adjacent to a canonical PAM sequence (NGG). In some embodiments, the 3' end of the target sequence is not immediately adjacent to a canonical PAM sequence (NGG). In some embodiments, the 3' end of the target sequence is immediately adjacent to an AGC, GAG, TTT, GTG, or CAA sequence.

[00348] In some embodiments, the target DNA sequence comprises a sequence

associated with a disease or disorder. In some embodiments, the target DNA sequence comprises a point mutation associated with a disease or disorder. In some embodiments, the activity of the fusion protein (*e.g.*, comprising an adenosine deaminase and a Cas9 domain), or the complex, results in a correction of the point mutation. In some embodiments, the target DNA sequence comprises a G→A point mutation associated with a disease or disorder, and wherein the deamination of the mutant A base results in a sequence that is not associated with a disease or disorder. In some embodiments, the target DNA sequence encodes a protein, and the point mutation is in a codon and results in a change in the amino acid encoded by the mutant codon as compared to the wild-type codon. In some embodiments, the deamination of the mutant A results in a change of the amino acid encoded by the mutant codon. In some embodiments, the deamination of the mutant A results in the codon encoding the wild-type amino acid. In some embodiments, the contacting is *in vivo* in a subject. In some embodiments, the subject has or has been diagnosed with a disease or disorder. In some embodiments, the disease or disorder is phenylketonuria, von Willebrand disease (vWD), a neoplastic disease associated with a mutant PTEN or BRCA1, or Li-Fraumeni syndrome. A list of exemplary diseases and disorders that may be treated using the nucleobase editors provided herein is shown in Table 1. Table 1 includes the target gene, the mutation to be corrected, the related disease and the nucleotide sequence of the associated protospacer and PAM.

Table 1 – List of exemplary diseases that may be treated using the nucleobase editors provided herein. The A to be edited in the protospacer is indicated by underlining and the PAM is indicated in bold.

Target Gene	Mutation	ATCC Cell Line	Disease	Protospacer and PAM
PTEN	Cys136Tyr	HTB-128	Cancer Predisposition	TATATGCATATTTATTACATCGG (SEQ ID NO: 85)
PTEN	Arg233Ter	HTB-13	Cancer Predisposition	CCGTCATGTGGGTCCTGAATTGG (SEQ ID NO: 86)
TP53	Glu258Lys	HTB-65	Cancer Predisposition	ACACTGAAAGACTCCAGGTCAGG (SEQ ID NO: 87)
BRCA1	Gly1738Arg	NA	Cancer Predisposition	GTCAGAAGAGATGTGGTCAATGG (SEQ ID NO: 88)
BRCA1	4097-1G>A	NA	Cancer Predisposition	TTTAAAGTGAAGCAGCATCTGGG (SEQ ID NO: 89); ATTTAAAGTGAAGCAGCATCTGG (SEQ ID NO: 90)
PAH	Thr380Met	NA	Phenylketonuria	ACTCCATGACAGTGTAATTTTGG (SEQ ID NO: 91)
VWF	Ser1285Phe	NA	von Willebrand (Hemophilia)	GCCTGGAGAAGCCATCCAGCAGG (SEQ ID NO: 92)
VWF	Arg2535Ter	NA	von Willebrand (Hemophilia)	CTCAGACACACTCATTGATGAGG (SEQ ID NO: 93)
TP53	Arg175His	HCC1395	Li-Fraumeni syndrome	GAGGCACTGCCCCCACCATGAGCG (SEQ ID NO: 94)

[00349] Some embodiments provide methods for using the DNA editing fusion proteins provided herein. In some embodiments, the fusion protein is used to introduce a

point mutation into a nucleic acid by deaminating a target nucleobase, *e.g.*, an A residue. In some embodiments, the deamination of the target nucleobase results in the correction of a genetic defect, *e.g.*, in the correction of a point mutation that leads to a loss of function in a gene product. In some embodiments, the genetic defect is associated with a disease or disorder, *e.g.*, a lysosomal storage disorder or a metabolic disease, such as, for example, type I diabetes. In some embodiments, the methods provided herein are used to introduce a deactivating point mutation into a gene or allele that encodes a gene product that is associated with a disease or disorder. For example, in some embodiments, methods are provided herein that employ a DNA editing fusion protein to introduce a deactivating point mutation into an oncogene (*e.g.*, in the treatment of a proliferative disease). A deactivating mutation may, in some embodiments, generate a premature stop codon in a coding sequence, which results in the expression of a truncated gene product, *e.g.*, a truncated protein lacking the function of the full-length protein.

[00350] In some embodiments, the purpose of the methods provided herein is to restore the function of a dysfunctional gene via genome editing. The nucleobase editing proteins provided herein can be validated for gene editing-based human therapeutics *in vitro*, *e.g.*, by correcting a disease-associated mutation in human cell culture. It will be understood by the skilled artisan that the nucleobase editing proteins provided herein, *e.g.*, the fusion proteins comprising a nucleic acid programmable DNA binding protein (*e.g.*, Cas9) and an adenosine deaminase domain can be used to correct any single point G to A or C to T mutation. In the first case, deamination of the mutant A to I corrects the mutation, and in the latter case, deamination of the A that is base-paired with the mutant T, followed by a round of replication, corrects the mutation. Exemplary point mutations that can be corrected are listed in Tables 1 and 2.

[00351] The successful correction of point mutations in disease-associated genes and alleles opens up new strategies for gene correction with applications in therapeutics and basic research. Site-specific single-base modification systems like the disclosed fusions of a nucleic acid programmable DNA binding protein and an adenosine deaminase domain also have applications in "reverse" gene therapy, where certain gene functions are purposely suppressed or abolished. In these cases, site-specifically mutating residues that lead to inactivating mutations in a protein, or mutations that inhibit function of the protein can be used to abolish or inhibit protein function *in vitro*, *ex vivo*, or *in vivo*.

[00352] The instant disclosure provides methods for the treatment of a subject diagnosed with a disease associated with or caused by a point mutation that can be corrected

by a DNA editing fusion protein provided herein. For example, in some embodiments, a method is provided that comprises administering to a subject having such a disease, *e.g.*, a cancer associated with a point mutation as described above, an effective amount of an adenosine deaminase fusion protein that corrects the point mutation or introduces a deactivating mutation into a disease-associated gene. In some embodiments, the disease is a proliferative disease. In some embodiments, the disease is a genetic disease. In some embodiments, the disease is a neoplastic disease. In some embodiments, the disease is a metabolic disease. In some embodiments, the disease is a lysosomal storage disease. Other diseases that can be treated by correcting a point mutation or introducing a deactivating mutation into a disease-associated gene will be known to those of skill in the art, and the disclosure is not limited in this respect.

[00353] The instant disclosure provides methods for the treatment of additional diseases or disorders, *e.g.*, diseases or disorders that are associated or caused by a point mutation that can be corrected by deaminase-mediated gene editing. Some such diseases are described herein, and additional suitable diseases that can be treated with the strategies and fusion proteins provided herein will be apparent to those of skill in the art based on the instant disclosure. Exemplary suitable diseases and disorders are listed below. It will be understood that the numbering of the specific positions or residues in the respective sequences depends on the particular protein and numbering scheme used. Numbering might be different, *e.g.*, in precursors of a mature protein and the mature protein itself, and differences in sequences from species to species may affect numbering. One of skill in the art will be able to identify the respective residue in any homologous protein and in the respective encoding nucleic acid by methods well known in the art, *e.g.*, by sequence alignment and determination of homologous residues. Exemplary suitable diseases and disorders include, without limitation: 2-methyl-3-hydroxybutyric aciduria; 3 beta-Hydroxysteroid dehydrogenase deficiency; 3-Methylglutaconic aciduria; 3-Oxo-5 alpha-steroid delta 4-dehydrogenase deficiency; 46,XY sex reversal, type 1, 3, and 5; 5-Oxoprolinase deficiency; 6-pyruvoyl-tetrahydropterin synthase deficiency; Aarskog syndrome; Aase syndrome; Achondrogenesis type 2; Achromatopsia 2 and 7; Acquired long QT syndrome; Acrocallosal syndrome, Schinzel type; Acrocapitofemoral dysplasia; Acrodysostosis 2, with or without hormone resistance; Acroerythrokeratoderma; Acromicric dysplasia; Acth-independent macronodular adrenal hyperplasia 2; Activated PI3K-delta syndrome; Acute intermittent porphyria; deficiency of Acyl-CoA dehydrogenase family, member 9; Adams-Oliver syndrome 5 and 6; Adenine phosphoribosyltransferase deficiency;

Adenylate kinase deficiency; hemolytic anemia due to Adenylosuccinate lyase deficiency; Adolescent nephronophthisis; Renal-hepatic-pancreatic dysplasia; Meckel syndrome type 7; Adrenoleukodystrophy; Adult junctional epidermolysis bullosa; Epidermolysis bullosa, junctional, localisata variant; Adult neuronal ceroid lipofuscinosis; Adult neuronal ceroid lipofuscinosis; Adult onset ataxia with oculomotor apraxia; ADULT syndrome; Afibrinogenemia and congenital Afibrinogenemia; autosomal recessive Agammaglobulinemia 2; Age-related macular degeneration 3, 6, 11, and 12; Aicardi Goutieres syndromes 1, 4, and 5; Chilbain lupus 1; Alagille syndromes 1 and 2; Alexander disease; Alkaptonuria; Allan-Herndon-Dudley syndrome; Alopecia universalis congenital; Alpers encephalopathy; Alpha-1-antitrypsin deficiency; autosomal dominant, autosomal recessive, and X-linked recessive Alport syndromes; Alzheimer disease, familial, 3, with spastic paraparesis and apraxia; Alzheimer disease, types, 1, 3, and 4; hypocalcification type and hypomaturation type, IIA1 Amelogenesis imperfecta; Aminoacylase 1 deficiency; Amish infantile epilepsy syndrome; Amyloidogenic transthyretin amyloidosis; Amyloid Cardiomyopathy, Transthyretin-related; Cardiomyopathy; Amyotrophic lateral sclerosis types 1, 6, 15 (with or without frontotemporal dementia), 22 (with or without frontotemporal dementia), and 10; Frontotemporal dementia with TDP43 inclusions, TARDBP-related; Andermann syndrome; Andersen Tawil syndrome; Congenital long QT syndrome; Anemia, nonspherocytic hemolytic, due to G6PD deficiency; Angelman syndrome; Severe neonatal-onset encephalopathy with microcephaly; susceptibility to Autism, X-linked 3; Angiopathy, hereditary, with nephropathy, aneurysms, and muscle cramps; Angiotensin i-converting enzyme, benign serum increase; Aniridia, cerebellar ataxia, and mental retardation; Anonychia; Antithrombin III deficiency; Antley-Bixler syndrome with genital anomalies and disordered steroidogenesis; Aortic aneurysm, familial thoracic 4, 6, and 9; Thoracic aortic aneurysms and aortic dissections; Multisystemic smooth muscle dysfunction syndrome; Moyamoya disease 5; Aplastic anemia; Apparent mineralocorticoid excess; Arginase deficiency; Argininosuccinate lyase deficiency; Aromatase deficiency; Arrhythmogenic right ventricular cardiomyopathy types 5, 8, and 10; Primary familial hypertrophic cardiomyopathy; Arthrogryposis multiplex congenita, distal, X-linked; Arthrogryposis renal dysfunction cholestasis syndrome; Arthrogryposis, renal dysfunction, and cholestasis 2; Asparagine synthetase deficiency; Abnormality of neuronal migration; Ataxia with vitamin E deficiency; Ataxia, sensory, autosomal dominant; Ataxia-telangiectasia syndrome; Hereditary cancer-predisposing syndrome; Atransferrinemia; Atrial fibrillation, familial, 11, 12, 13, and 16; Atrial septal defects 2, 4, and 7 (with or without atrioventricular conduction defects);

Atrial standstill 2; Atrioventricular septal defect 4; Atrophia bulborum hereditaria; ATR-X syndrome; Auriculocondylar syndrome 2; Autoimmune disease, multisystem, infantile-onset; Autoimmune lymphoproliferative syndrome, type 1a; Autosomal dominant hypohidrotic ectodermal dysplasia; Autosomal dominant progressive external ophthalmoplegia with mitochondrial DNA deletions 1 and 3; Autosomal dominant torsion dystonia 4; Autosomal recessive centronuclear myopathy; Autosomal recessive congenital ichthyosis 1, 2, 3, 4A, and 4B; Autosomal recessive cutis laxa type IA and 1B; Autosomal recessive hypohidrotic ectodermal dysplasia syndrome; Ectodermal dysplasia 11b; hypohidrotic/hair/tooth type, autosomal recessive; Autosomal recessive hypophosphatemic bone disease; Axenfeld-Rieger syndrome type 3; Bainbridge-Ropers syndrome; Bannayan-Riley-Ruvalcaba syndrome; PTEN hamartoma tumor syndrome; Baraitser-Winter syndromes 1 and 2; Barakat syndrome; Bardet-Biedl syndromes 1, 11, 16, and 19; Bare lymphocyte syndrome type 2, complementation group E; Bartter syndrome antenatal type 2; Bartter syndrome types 3, 3 with hypocalciuria , and 4; Basal ganglia calcification, idiopathic, 4; Beaded hair; Benign familial hematuria; Benign familial neonatal seizures 1 and 2; Seizures, benign familial neonatal, 1, and/or myokymia; Seizures, Early infantile epileptic encephalopathy 7; Benign familial neonatal-infantile seizures; Benign hereditary chorea; Benign scapuloperoneal muscular dystrophy with cardiomyopathy; Bernard-Soulier syndrome, types A1 and A2 (autosomal dominant); Bestrophinopathy, autosomal recessive; beta Thalassemia; Bethlem myopathy and Bethlem myopathy 2; Bietti crystalline corneoretinal dystrophy; Bile acid synthesis defect, congenital, 2; Biotinidase deficiency; Birk Barel mental retardation dysmorphism syndrome; Blepharophimosis, ptosis, and epicanthus inversus; Bloom syndrome; Borjeson-Forssman-Lehmann syndrome; Boucher Neuhauser syndrome; Brachydactyly types A1 and A2; Brachydactyly with hypertension; Brain small vessel disease with hemorrhage; Branched-chain ketoacid dehydrogenase kinase deficiency; Branchiootic syndromes 2 and 3; Breast cancer, early-onset; Breast-ovarian cancer, familial 1, 2, and 4; Brittle cornea syndrome 2; Brody myopathy; Bronchiectasis with or without elevated sweat chloride 3; Brown-Vialetto-Van laere syndrome and Brown-Vialetto-Van Laere syndrome 2; Brugada syndrome; Brugada syndrome 1; Ventricular fibrillation; Paroxysmal familial ventricular fibrillation; Brugada syndrome and Brugada syndrome 4; Long QT syndrome; Sudden cardiac death; Bull eye macular dystrophy; Stargardt disease 4; Cone-rod dystrophy 12; Bullous ichthyosiform erythroderma; Burn-Mckeown syndrome; Candidiasis, familial, 2, 5, 6, and 8; Carbohydrate-deficient glycoprotein syndrome type I and II; Carbonic anhydrase VA deficiency, hyperammonemia due to; Carcinoma of colon;

Cardiac arrhythmia; Long QT syndrome, LQT1 subtype; Cardioencephalomyopathy, fatal infantile, due to cytochrome c oxidase deficiency; Cardiofaciocutaneous syndrome; Cardiomyopathy; Danon disease; Hypertrophic cardiomyopathy; Left ventricular noncompaction cardiomyopathy; Carnevale syndrome; Carney complex, type 1; Carnitine acylcarnitine translocase deficiency; Carnitine palmitoyltransferase I , II, II (late onset), and II (infantile) deficiency; Cataract 1, 4, autosomal dominant, autosomal dominant, multiple types, with microcornea, coppock-like, juvenile, with microcornea and glucosuria, and nuclear diffuse nonprogressive; Catecholaminergic polymorphic ventricular tachycardia; Caudal regression syndrome; Cd8 deficiency, familial; Central core disease; Centromeric instability of chromosomes 1,9 and 16 and immunodeficiency; Cerebellar ataxia infantile with progressive external ophthalmoplegi and Cerebellar ataxia, mental retardation, and dysequilibrium syndrome 2; Cerebral amyloid angiopathy, APP-related; Cerebral autosomal dominant and recessive arteriopathy with subcortical infarcts and leukoencephalopathy; Cerebral cavernous malformations 2; Cerebrooculofacioskeletal syndrome 2; Cerebro-oculo-facio-skeletal syndrome; Cerebroretinal microangiopathy with calcifications and cysts; Ceroid lipofuscinosis neuronal 2, 6, 7, and 10; Ch\xc3\xa9diak-Higashi syndrome , Chediak-Higashi syndrome, adult type; Charcot-Marie-Tooth disease types 1B, 2B2, 2C, 2F, 2I, 2U (axonal), 1C (demyelinating), dominant intermediate C, recessive intermediate A, 2A2, 4C, 4D, 4H, IF, IVF, and X; Scapuloperoneal spinal muscular atrophy; Distal spinal muscular atrophy, congenital nonprogressive; Spinal muscular atrophy, distal, autosomal recessive, 5; CHARGE association; Childhood hypophosphatasia; Adult hypophosphatasia; Cholecystitis; Progressive familial intrahepatic cholestasis 3; Cholestasis, intrahepatic, of pregnancy 3; Cholestanol storage disease; Cholesterol monooxygenase (side-chain cleaving) deficiency; Chondrodysplasia Blomstrand type; Chondrodysplasia punctata 1, X-linked recessive and 2 X-linked dominant; CHOPS syndrome; Chronic granulomatous disease, autosomal recessive cytochrome b-positive, types 1 and 2; Chudley-McCullough syndrome; Ciliary dyskinesia, primary, 7, 11, 15, 20 and 22; Citrullinemia type I; Citrullinemia type I and II; Cleidocranial dysostosis; C-like syndrome; Cockayne syndrome type A, ; Coenzyme Q10 deficiency, primary 1, 4, and 7; Coffin Siris/Intellectual Disability; Coffin-Lowry syndrome; Cohen syndrome, ; Cold-induced sweating syndrome 1; COLE-CARPENTER SYNDROME 2; Combined cellular and humoral immune defects with granulomas; Combined d-2- and l-2-hydroxyglutaric aciduria; Combined malonic and methylmalonic aciduria; Combined oxidative phosphorylation deficiencies 1, 3, 4, 12, 15, and 25; Combined partial and complete 17-alpha-hydroxylase/17,20-lyase deficiency; Common variable immunodeficiency 9;

Complement component 4, partial deficiency of, due to dysfunctional c1 inhibitor; Complement factor B deficiency; Cone monochromatism; Cone-rod dystrophy 2 and 6; Cone-rod dystrophy amelogenesis imperfecta; Congenital adrenal hyperplasia and Congenital adrenal hypoplasia, X-linked; Congenital amegakaryocytic thrombocytopenia; Congenital aniridia; Congenital central hypoventilation; Hirschsprung disease 3; Congenital contractural arachnodactyly; Congenital contractures of the limbs and face, hypotonia, and developmental delay; Congenital disorder of glycosylation types 1B, 1D, 1G, 1H, 1J, 1K, 1N, 1P, 2C, 2J, 2K, IIm; Congenital dyserythropoietic anemia, type I and II; Congenital ectodermal dysplasia of face; Congenital erythropoietic porphyria; Congenital generalized lipodystrophy type 2; Congenital heart disease, multiple types, 2; Congenital heart disease; Interrupted aortic arch; Congenital lipomatous overgrowth, vascular malformations, and epidermal nevi; Non-small cell lung cancer; Neoplasm of ovary; Cardiac conduction defect, nonspecific; Congenital microvillous atrophy; Congenital muscular dystrophy; Congenital muscular dystrophy due to partial LAMA2 deficiency; Congenital muscular dystrophy-dystroglycanopathy with brain and eye anomalies, types A2, A7, A8, A11, and A14; Congenital muscular dystrophy-dystroglycanopathy with mental retardation, types B2, B3, B5, and B15; Congenital muscular dystrophy-dystroglycanopathy without mental retardation, type B5; Congenital muscular hypertrophy-cerebral syndrome; Congenital myasthenic syndrome, acetazolamide-responsive; Congenital myopathy with fiber type disproportion; Congenital ocular coloboma; Congenital stationary night blindness, type 1A, 1B, 1C, 1E, 1F, and 2A; Coproporphyria; Cornea plana 2; Corneal dystrophy, Fuchs endothelial, 4; Corneal endothelial dystrophy type 2; Corneal fragility keratoglobus, blue sclerae and joint hypermobility; Cornelia de Lange syndromes 1 and 5; Coronary artery disease, autosomal dominant 2; Coronary heart disease; Hyperalphalipoproteinemia 2; Cortical dysplasia, complex, with other brain malformations 5 and 6; Cortical malformations, occipital; Corticosteroid-binding globulin deficiency; Corticosterone methyloxidase type 2 deficiency; Costello syndrome; Cowden syndrome 1; Coxa plana; Craniodiaphyseal dysplasia, autosomal dominant; Craniosynostosis 1 and 4; Craniosynostosis and dental anomalies; Creatine deficiency, X-linked; Crouzon syndrome; Cryptophthalmos syndrome; Cryptorchidism, unilateral or bilateral; Cushing symphalangism; Cutaneous malignant melanoma 1; Cutis laxa with osteodystrophy and with severe pulmonary, gastrointestinal, and urinary abnormalities; Cyanosis, transient neonatal and atypical nephropathic; Cystic fibrosis; Cystinuria; Cytochrome c oxidase i deficiency; Cytochrome-c oxidase deficiency ; D-2-hydroxyglutaric aciduria 2; Darier disease, segmental; Deafness with labyrinthine aplasia microtia and microdontia (LAMM); Deafness,

autosomal dominant 3a, 4, 12, 13, 15, autosomal dominant nonsyndromic sensorineural 17, 20, and 65; Deafness, autosomal recessive 1A, 2, 3, 6, 8, 9, 12, 15, 16, 18b, 22, 28, 31, 44, 49, 63, 77, 86, and 89; Deafness, cochlear, with myopia and intellectual impairment, without vestibular involvement, autosomal dominant, X-linked 2; Deficiency of 2-methylbutyryl-CoA dehydrogenase; Deficiency of 3-hydroxyacyl-CoA dehydrogenase; Deficiency of alpha-mannosidase; Deficiency of aromatic-L-amino-acid decarboxylase; Deficiency of bisphosphoglycerate mutase; Deficiency of butyryl-CoA dehydrogenase; Deficiency of ferroxidase; Deficiency of galactokinase; Deficiency of guanidinoacetate methyltransferase; Deficiency of hyaluronoglucosaminidase; Deficiency of ribose-5-phosphate isomerase; Deficiency of steroid 11-beta-monooxygenase; Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase; Deficiency of xanthine oxidase; Dejerine-Sottas disease; Charcot-Marie-Tooth disease, types ID and IVF; Dejerine-Sottas syndrome, autosomal dominant; Dendritic cell, monocyte, B lymphocyte, and natural killer lymphocyte deficiency; Desbuquois dysplasia 2; Desbuquois syndrome; DFNA 2 Nonsyndromic Hearing Loss; Diabetes mellitus and insipidus with optic atrophy and deafness; Diabetes mellitus, type 2, and insulin-dependent, 20; Diamond-Blackfan anemia 1, 5, 8, and 10; Diarrhea 3 (secretory sodium, congenital, syndromic) and 5 (with tufting enteropathy, congenital); Dicarboxylic aminoaciduria; Diffuse palmoplantar keratoderma, Bothnian type; Digitorenocerebral syndrome; Dihydropteridine reductase deficiency; Dilated cardiomyopathy 1A, 1AA, 1C, 1G, 1BB, 1DD, 1FF, 1HH, 1I, 1KK, 1N, 1S, 1Y, and 3B; Left ventricular noncompaction 3; Disordered steroidogenesis due to cytochrome p450 oxidoreductase deficiency; Distal arthrogryposis type 2B; Distal hereditary motor neuronopathy type 2B; Distal myopathy Markesbery-Griggs type; Distal spinal muscular atrophy, X-linked 3; Distichiasis-lymphedema syndrome; Dominant dystrophic epidermolysis bullosa with absence of skin; Dominant hereditary optic atrophy; Donnai Barrow syndrome; Dopamine beta hydroxylase deficiency; Dopamine receptor d2, reduced brain density of; Dowling-degos disease 4; Doyne honeycomb retinal dystrophy; Malattia leventinese; Duane syndrome type 2; Dubin-Johnson syndrome; Duchenne muscular dystrophy; Becker muscular dystrophy; Dysfibrinogenemia; Dyskeratosis congenita autosomal dominant and autosomal dominant, 3; Dyskeratosis congenita, autosomal recessive, 1, 3, 4, and 5; Dyskeratosis congenita X-linked; Dyskinesia, familial, with facial myokymia; Dysplasminogenemia; Dystonia 2 (torsion, autosomal recessive), 3 (torsion, X-linked), 5 (Dopa-responsive type), 10, 12, 16, 25, 26 (Myoclonic); Seizures, benign familial infantile, 2; Early infantile epileptic encephalopathy 2, 4, 7, 9, 10, 11, 13, and 14; Atypical Rett syndrome; Early T cell progenitor acute lymphoblastic

leukemia; Ectodermal dysplasia skin fragility syndrome; Ectodermal dysplasia-syndactyly syndrome 1; Ectopia lentis, isolated autosomal recessive and dominant; Ectrodactyly, ectodermal dysplasia, and cleft lip/palate syndrome 3; Ehlers-Danlos syndrome type 7 (autosomal recessive), classic type, type 2 (progeroid), hydroxylysine-deficient, type 4, type 4 variant, and due to tenascin-X deficiency; Eichsfeld type congenital muscular dystrophy; Endocrine-cerebroosteodysplasia; Enhanced s-cone syndrome; Enlarged vestibular aqueduct syndrome; Enterokinase deficiency; Epidermodysplasia verruciformis; Epidermolysa bullosa simplex and limb girdle muscular dystrophy, simplex with mottled pigmentation, simplex with pyloric atresia, simplex, autosomal recessive, and with pyloric atresia; Epidermolytic palmoplantar keratoderma; Familial febrile seizures 8; Epilepsy, childhood absence 2, 12 (idiopathic generalized, susceptibility to) 5 (nocturnal frontal lobe), nocturnal frontal lobe type 1, partial, with variable foci, progressive myoclonic 3, and X-linked, with variable learning disabilities and behavior disorders; Epileptic encephalopathy, childhood-onset, early infantile, 1, 19, 23, 25, 30, and 32; Epiphyseal dysplasia, multiple, with myopia and conductive deafness; Episodic ataxia type 2; Episodic pain syndrome, familial, 3; Epstein syndrome; Fechtner syndrome; Erythropoietic protoporphyria; Estrogen resistance; Exudative vitreoretinopathy 6; Fabry disease and Fabry disease, cardiac variant; Factor H, VII, X, v and factor viii, combined deficiency of 2, xiii, a subunit, deficiency; Familial adenomatous polyposis 1 and 3; Familial amyloid nephropathy with urticaria and deafness; Familial cold urticarial; Familial aplasia of the vermis; Familial benign pemphigus; Familial cancer of breast; Breast cancer, susceptibility to; Osteosarcoma; Pancreatic cancer 3; Familial cardiomyopathy; Familial cold autoinflammatory syndrome 2; Familial colorectal cancer; Familial exudative vitreoretinopathy, X-linked; Familial hemiplegic migraine types 1 and 2; Familial hypercholesterolemia; Familial hypertrophic cardiomyopathy 1, 2, 3, 4, 7, 10, 23 and 24; Familial hypokalemia-hypomagnesemia; Familial hypoplastic, glomerulocystic kidney; Familial infantile myasthenia; Familial juvenile gout; Familial Mediterranean fever and Familial mediterranean fever, autosomal dominant; Familial porencephaly; Familial porphyria cutanea tarda; Familial pulmonary capillary hemangiomatosis; Familial renal glucosuria; Familial renal hypouricemia; Familial restrictive cardiomyopathy 1; Familial type 1 and 3 hyperlipoproteinemia; Fanconi anemia, complementation group E, I, N, and O; Fanconi-Bickel syndrome; Favism, susceptibility to; Febrile seizures, familial, 11; Feingold syndrome 1; Fetal hemoglobin quantitative trait locus 1; FG syndrome and FG syndrome 4; Fibrosis of extraocular muscles, congenital, 1, 2, 3a (with or without extraocular involvement), 3b; Fish-eye disease; Fleck corneal dystrophy; Floating-Harbor syndrome;

Focal epilepsy with speech disorder with or without mental retardation; Focal segmental glomerulosclerosis 5; Forebrain defects; Frank Ter Haar syndrome; Borrone Di Rocco Crovato syndrome; Frasier syndrome; Wilms tumor 1; Freeman-Sheldon syndrome; Frontometaphyseal dysplasia 1and 3; Frontotemporal dementia; Frontotemporal dementia and/or amyotrophic lateral sclerosis 3 and 4; Frontotemporal Dementia Chromosome 3-Linked and Frontotemporal dementia ubiquitin-positive; Fructose-biphosphatase deficiency; Fuhrmann syndrome; Gamma-aminobutyric acid transaminase deficiency; Gamstorp-Wohlfart syndrome; Gaucher disease type 1 and Subacute neuronopathic; Gaze palsy, familial horizontal, with progressive scoliosis; Generalized dominant dystrophic epidermolysis bullosa; Generalized epilepsy with febrile seizures plus 3, type 1, type 2; Epileptic encephalopathy Lennox-Gastaut type; Giant axonal neuropathy; Glanzmann thrombasthenia; Glaucoma 1, open angle, e, F, and G; Glaucoma 3, primary congenital, d; Glaucoma, congenital and Glaucoma, congenital, Coloboma; Glaucoma, primary open angle, juvenile-onset; Glioma susceptibility 1; Glucose transporter type 1 deficiency syndrome; Glucose-6-phosphate transport defect; GLUT1 deficiency syndrome 2; Epilepsy, idiopathic generalized, susceptibility to, 12; Glutamate formiminotransferase deficiency; Glutaric acidemia IIA and IIB; Glutaric aciduria, type 1; Gluthathione synthetase deficiency; Glycogen storage disease 0 (muscle), II (adult form), IXa2, IXc, type 1A; type II, type IV, IV (combined hepatic and myopathic), type V, and type VI; Goldmann-Favre syndrome; Gordon syndrome; Gorlin syndrome; Holoprosencephaly sequence; Holoprosencephaly 7; Granulomatous disease, chronic, X-linked, variant; Granulosa cell tumor of the ovary; Gray platelet syndrome; Griscelli syndrome type 3; Groenouw corneal dystrophy type I; Growth and mental retardation, mandibulofacial dysostosis, microcephaly, and cleft palate; Growth hormone deficiency with pituitary anomalies; Growth hormone insensitivity with immunodeficiency; GTP cyclohydrolase I deficiency; Hajdu-Cheney syndrome; Hand foot uterus syndrome; Hearing impairment; Hemangioma, capillary infantile; Hematologic neoplasm; Hemochromatosis type 1, 2B, and 3; Microvascular complications of diabetes 7; Transferrin serum level quantitative trait locus 2; Hemoglobin H disease, nondeletional; Hemolytic anemia, nonspherocytic, due to glucose phosphate isomerase deficiency; Hemophagocytic lymphohistiocytosis, familial, 2; Hemophagocytic lymphohistiocytosis, familial, 3; Heparin cofactor II deficiency; Hereditary acrodermatitis enteropathica; Hereditary breast and ovarian cancer syndrome; Ataxia-telangiectasia-like disorder; Hereditary diffuse gastric cancer; Hereditary diffuse leukoencephalopathy with spheroids; Hereditary factors II, IX, VIII deficiency disease; Hereditary hemorrhagic telangiectasia type

2; Hereditary insensitivity to pain with anhidrosis; Hereditary lymphedema type I; Hereditary motor and sensory neuropathy with optic atrophy; Hereditary myopathy with early respiratory failure; Hereditary neuralgic amyotrophy; Hereditary Nonpolyposis Colorectal Neoplasms; Lynch syndrome I and II; Hereditary pancreatitis; Pancreatitis, chronic, susceptibility to; Hereditary sensory and autonomic neuropathy type IIB amd IIA; Hereditary sideroblastic anemia; Hermansky-Pudlak syndrome 1, 3, 4, and 6; Heterotaxy, visceral, 2, 4, and 6, autosomal; Heterotaxy, visceral, X-linked; Heterotopia; Histiocytic medullary reticulosis; Histiocytosis-lymphadenopathy plus syndrome; Holocarboxylase synthetase deficiency; Holoprosencephaly 2, 3,7, and 9; Holt-Oram syndrome; Homocysteinemia due to MTHFR deficiency, CBS deficiency, and Homocystinuria, pyridoxine-responsive; Homocystinuria-Megaloblastic anemia due to defect in cobalamin metabolism, cblE complementation type; Howel-Evans syndrome; Hurler syndrome; Hutchinson-Gilford syndrome; Hydrocephalus; Hyperammonemia, type III; Hypercholesterolaemia and Hypercholesterolemia, autosomal recessive; Hyperekplexia 2 and Hyperekplexia hereditary; Hyperferritinemia cataract syndrome; Hyperglycinuria; Hyperimmunoglobulin D with periodic fever; Mevalonic aciduria; Hyperimmunoglobulin E syndrome; Hyperinsulinemic hypoglycemia familial 3, 4, and 5; Hyperinsulinism-hyperammonemia syndrome; Hyperlysinemia; Hypermanganesemia with dystonia, polycythemia and cirrhosis; Hyperornithinemia-hyperammonemia-homocitrullinuria syndrome; Hyperparathyroidism 1 and 2; Hyperparathyroidism, neonatal severe; Hyperphenylalaninemia, bh4-deficient, a, due to partial pts deficiency, BH4-deficient, D, and non-pku; Hyperphosphatasia with mental retardation syndrome 2, 3, and 4; Hypertrichotic osteochondrodysplasia; Hypobetalipoproteinemia, familial, associated with apob32; Hypocalcemia, autosomal dominant 1; Hypocalciuric hypercalcemia, familial, types 1 and 3; Hypochondrogenesis; Hypochromic microcytic anemia with iron overload; Hypoglycemia with deficiency of glycogen synthetase in the liver; Hypogonadotropic hypogonadism 11 with or without anosmia; Hypohidrotic ectodermal dysplasia with immune deficiency; Hypohidrotic X-linked ectodermal dysplasia; Hypokalemic periodic paralysis 1 and 2; Hypomagnesemia 1, intestinal; Hypomagnesemia, seizures, and mental retardation; Hypomyelinating leukodystrophy 7; Hypoplastic left heart syndrome; Atrioventricular septal defect and common atrioventricular junction; Hypospadias 1 and 2, X-linked; Hypothyroidism, congenital, nongoitrous, 1; Hypotrichosis 8 and 12; Hypotrichosis-lymphedema-telangiectasia syndrome; I blood group system; Ichthyosis bullosa of Siemens; Ichthyosis exfoliativa; Ichthyosis prematurity syndrome; Idiopathic basal ganglia calcification 5;

Idiopathic fibrosing alveolitis, chronic form; Dyskeratosis congenita, autosomal dominant, 2 and 5; Idiopathic hypercalcemia of infancy; Immune dysfunction with T-cell inactivation due to calcium entry defect 2; Immunodeficiency 15, 16, 19, 30, 31C, 38, 40, 8, due to defect in cd3-zeta, with hyper IgM type 1 and 2, and X-Linked, with magnesium defect, Epstein-Barr virus infection, and neoplasia; Immunodeficiency-centromeric instability-facial anomalies syndrome 2; Inclusion body myopathy 2 and 3; Nonaka myopathy; Infantile convulsions and paroxysmal choreoathetosis, familial; Infantile cortical hyperostosis; Infantile GM1 gangliosidosis; Infantile hypophosphatasia; Infantile nephronophthisis; Infantile nystagmus, X-linked; Infantile Parkinsonism-dystonia; Infertility associated with multi-tailed spermatozoa and excessive DNA; Insulin resistance; Insulin-resistant diabetes mellitus and acanthosis nigricans; Insulin-dependent diabetes mellitus secretory diarrhea syndrome; Interstitial nephritis, karyomegalic; Intrauterine growth retardation, metaphyseal dysplasia, adrenal hypoplasia congenita, and genital anomalies; Iodotyrosyl coupling defect; IRAK4 deficiency; Iridogoniodysgenesis dominant type and type 1; Iron accumulation in brain; Ischiopatellar dysplasia; Islet cell hyperplasia; Isolated 17,20-lyase deficiency; Isolated lutropin deficiency; Isovaleryl-CoA dehydrogenase deficiency; Jankovic Rivera syndrome; Jervell and Lange-Nielsen syndrome 2; Joubert syndrome 1, 6, 7, 9/15 (digenic), 14, 16, and 17, and Orofaciodigital syndrome xiv; Junctional epidermolysis bullosa gravis of Herlitz; Juvenile GM>1< gangliosidosis; Juvenile polyposis syndrome; Juvenile polyposis/hereditary hemorrhagic telangiectasia syndrome; Juvenile retinoschisis; Kabuki make-up syndrome; Kallmann syndrome 1, 2, and 6; Delayed puberty; Kanzaki disease; Karak syndrome; Kartagener syndrome; Kenny-Caffey syndrome type 2; Keppen-Lubinsky syndrome; Keratoconus 1; Keratosis follicularis; Keratosis palmoplantaris striata 1; Kindler syndrome; L-2-hydroxyglutaric aciduria; Larsen syndrome, dominant type; Lattice corneal dystrophy Type III; Leber amaurosis; Zellweger syndrome; Peroxisome biogenesis disorders; Zellweger syndrome spectrum; Leber congenital amaurosis 11, 12, 13, 16, 4, 7, and 9; Leber optic atrophy; Aminoglycoside-induced deafness; Deafness, nonsyndromic sensorineural, mitochondrial; Left ventricular noncompaction 5; Left-right axis malformations; Leigh disease; Mitochondrial short-chain Enoyl-CoA Hydratase 1 deficiency; Leigh syndrome due to mitochondrial complex I deficiency; Leiner disease; Leri Weill dyschondrosteosis; Lethal congenital contracture syndrome 6; Leukocyte adhesion deficiency type I and III; Leukodystrophy, Hypomyelinating, 11 and 6; Leukoencephalopathy with ataxia, with Brainstem and Spinal Cord Involvement and Lactate Elevation, with vanishing white matter, and progressive, with ovarian failure; Leukonychia totalis; Lewy body dementia;

Lichtenstein-Knorr Syndrome; Li-Fraumeni syndrome 1; Lig4 syndrome; Limb-girdle muscular dystrophy, type 1B, 2A, 2B, 2D, C1, C5, C9, C14; Congenital muscular dystrophy-dystroglycanopathy with brain and eye anomalies, type A14 and B14; Lipase deficiency combined; Lipid proteinosis; Lipodystrophy, familial partial, type 2 and 3; Lissencephaly 1, 2 (X-linked), 3, 6 (with microcephaly), X-linked; Subcortical laminar heterotopia, X-linked; Liver failure acute infantile; Loeys-Dietz syndrome 1, 2, 3; Long QT syndrome 1, 2, 2/9, 2/5, (digenic), 3, 5 and 5, acquired, susceptibility to; Lung cancer; Lymphedema, hereditary, id; Lymphedema, primary, with myelodysplasia; Lymphoproliferative syndrome 1, 1 (X-linked), and 2; Lysosomal acid lipase deficiency; Macrocephaly, macrosomia, facial dysmorphism syndrome; Macular dystrophy, vitelliform, adult-onset; Malignant hyperthermia susceptibility type 1; Malignant lymphoma, non-Hodgkin; Malignant melanoma; Malignant tumor of prostate; Mandibuloacral dysostosis; Mandibuloacral dysplasia with type A or B lipodystrophy, atypical; Mandibulofacial dysostosis, Treacher Collins type, autosomal recessive; Mannose-binding protein deficiency; Maple syrup urine disease type 1A and type 3; Marden Walker like syndrome; Marfan syndrome; Marinesco-Sj\xc3\xb6gren syndrome; Martsolf syndrome; Maturity-onset diabetes of the young, type 1, type 2, type 11, type 3, and type 9; May-Hegglin anomaly; MYH9 related disorders; Sebastian syndrome; McCune-Albright syndrome; Somatotroph adenoma; Sex cord-stromal tumor; Cushing syndrome; McKusick Kaufman syndrome; McLeod neuroacanthocytosis syndrome; Meckel-Gruber syndrome; Medium-chain acyl-coenzyme A dehydrogenase deficiency; Medulloblastoma; Megalencephalic leukoencephalopathy with subcortical cysts 1and 2a; Megalencephaly cutis marmorata telangiectatica congenital; PIK3CA Related Overgrowth Spectrum; Megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome 2; Megaloblastic anemia, thiamine-responsive, with diabetes mellitus and sensorineural deafness; Meier-Gorlin syndromes 1and 4; Melnick-Needles syndrome; Meningioma; Mental retardation, X-linked, 3, 21, 30, and 72; Mental retardation and microcephaly with pontine and cerebellar hypoplasia; Mental retardation X-linked syndromic 5; Mental retardation, anterior maxillary protrusion, and strabismus; Mental retardation, autosomal dominant 12, 13, 15, 24, 3, 30, 4, 5, 6,and 9; Mental retardation, autosomal recessive 15, 44, 46, and 5; Mental retardation, stereotypic movements, epilepsy, and/or cerebral malformations; Mental retardation, syndromic, Claes-Jensen type, X-linked; Mental retardation, X-linked, nonspecific, syndromic, Hedera type, and syndromic, wu type; Merosin deficient congenital muscular dystrophy; Metachromatic leukodystrophy juvenile, late infantile, and adult types; Metachromatic leukodystrophy; Metatrophic dysplasia; Methemoglobinemia types I and 2;

Methionine adenosyltransferase deficiency, autosomal dominant; Methylmalonic acidemia with homocystinuria, ; Methylmalonic aciduria cblB type, ; Methylmalonic aciduria due to methylmalonyl-CoA mutase deficiency; METHYLMALONIC ACIDURIA, mut(0) TYPE; Microcephalic osteodysplastic primordial dwarfism type 2; Microcephaly with or without chorioretinopathy, lymphedema, or mental retardation; Microcephaly, hiatal hernia and nephrotic syndrome; Microcephaly; Hypoplasia of the corpus callosum; Spastic paraplegia 50, autosomal recessive; Global developmental delay; CNS hypomyelination; Brain atrophy; Microcephaly, normal intelligence and immunodeficiency; Microcephaly-capillary malformation syndrome; Microcytic anemia; Microphthalmia syndromic 5, 7, and 9; Microphthalmia, isolated 3, 5, 6, 8, and with coloboma 6; Microspherophakia; Migraine, familial basilar; Miller syndrome; Minicore myopathy with external ophthalmoplegia; Myopathy, congenital with cores; Mitchell-Riley syndrome; mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase deficiency; Mitochondrial complex I, II, III, III (nuclear type 2, 4, or 8) deficiency; Mitochondrial DNA depletion syndrome 11, 12 (cardiomyopathic type), 2, 4B (MNGIE type), 8B (MNGIE type); Mitochondrial DNA-depletion syndrome 3 and 7, hepatocerebral types, and 13 (encephalomyopathic type); Mitochondrial phosphate carrier and pyruvate carrier deficiency; Mitochondrial trifunctional protein deficiency; Long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency; Miyoshi muscular dystrophy 1; Myopathy, distal, with anterior tibial onset; Mohr-Tranebjaerg syndrome; Molybdenum cofactor deficiency, complementation group A; Mowat-Wilson syndrome; Mucolipidosis III Gamma; Mucopolysaccharidosis type VI, type VI (severe), and type VII; Mucopolysaccharidosis, MPS-I-H/S, MPS-II, MPS-III-A, MPS-III-B, MPS-III-C, MPS-IV-A, MPS-IV-B; Retinitis Pigmentosa 73; Gangliosidosis GM1 type1 (with cardiac involvenment) 3; Multicentric osteolysis nephropathy; Multicentric osteolysis, nodulosis and arthropathy; Multiple congenital anomalies; Atrial septal defect 2; Multiple congenital anomalies-hypotonia-seizures syndrome 3; Multiple Cutaneous and Mucosal Venous Malformations; Multiple endocrine neoplasia, types 1and 4; Multiple epiphyseal dysplasia 5 or Dominant; Multiple gastrointestinal atresias; Multiple pterygium syndrome Escobar type; Multiple sulfatase deficiency; Multiple synostoses syndrome 3; Muscle AMP deaminase deficiency; Muscle eye brain disease; Muscular dystrophy, congenital, megaconial type; Myasthenia, familial infantile, 1; Myasthenic Syndrome, Congenital, 11, associated with acetylcholine receptor deficiency; Myasthenic Syndrome, Congenital, 17, 2A (slow-channel), 4B (fast-channel), and without tubular aggregates; Myeloperoxidase deficiency; MYH-associated polyposis; Endometrial carcinoma; Myocardial infarction 1; Myoclonic dystonia; Myoclonic-Atonic

Epilepsy; Myoclonus with epilepsy with ragged red fibers; Myofibrillar myopathy 1 and ZASP-related; Myoglobinuria, acute recurrent, autosomal recessive; Myoneural gastrointestinal encephalopathy syndrome; Cerebellar ataxia infantile with progressive external ophthalmoplegia; Mitochondrial DNA depletion syndrome 4B, MNGIE type; Myopathy, centronuclear, 1, congenital, with excess of muscle spindles, distal, 1, lactic acidosis, and sideroblastic anemia 1, mitochondrial progressive with congenital cataract, hearing loss, and developmental delay, and tubular aggregate, 2; Myopia 6; Myosclerosis, autosomal recessive; Myotonia congenital; Congenital myotonia, autosomal dominant and recessive forms; Nail-patella syndrome; Nance-Horan syndrome; Nanophthalmos 2; Navajo neurohepatopathy; Nemaline myopathy 3 and 9; Neonatal hypotonia; Intellectual disability; Seizures; Delayed speech and language development; Mental retardation, autosomal dominant 31; Neonatal intrahepatic cholestasis caused by citrin deficiency; Nephrogenic diabetes insipidus, Nephrogenic diabetes insipidus, X-linked; Nephrolithiasis/osteoporosis, hypophosphatemic, 2; Nephronophthisis 13, 15 and 4; Infertility; Cerebello-oculo-renal syndrome (nephronophthisis, oculomotor apraxia and cerebellar abnormalities); Nephrotic syndrome, type 3, type 5, with or without ocular abnormalities, type 7, and type 9; Nestor-Guillermo progeria syndrome; Neu-Laxova syndrome 1; Neurodegeneration with brain iron accumulation 4 and 6; Neuroferritinopathy; Neurofibromatosis, type 1and type 2; Neurofibrosarcoma; Neurohypophyseal diabetes insipidus; Neuropathy, Hereditary Sensory, Type IC; Neutral 1 amino acid transport defect; Neutral lipid storage disease with myopathy; Neutrophil immunodeficiency syndrome; Nicolaides-Baraitser syndrome; Niemann-Pick disease type C1, C2, type A, and type C1, adult form; Non-ketotic hyperglycinemia; Noonan syndrome 1 and 4, LEOPARD syndrome 1; Noonan syndrome-like disorder with or without juvenile myelomonocytic leukemia; Normokalemic periodic paralysis, potassium-sensitive; Norum disease; Epilepsy, Hearing Loss, And Mental Retardation Syndrome; Mental Retardation, X-Linked 102 and syndromic 13; Obesity; Ocular albinism, type I; Oculocutaneous albinism type 1B, type 3, and type 4; Oculodentodigital dysplasia; Odontohypophosphatasia; Odontotrichomelic syndrome; Oguchi disease; Oligodontia-colorectal cancer syndrome; Opitz G/BBB syndrome; Optic atrophy 9; Oral-facial-digital syndrome; Ornithine aminotransferase deficiency; Orofacial cleft 11 and 7, Cleft lip/palate-ectodermal dysplasia syndrome; Orstavik Lindemann Solberg syndrome; Osteoarthritis with mild chondrodysplasia; Osteochondritis dissecans; Osteogenesis imperfecta type 12, type 5, type 7, type 8, type I, type III, with normal sclerae, dominant form, recessive perinatal lethal; Osteopathia striata with cranial sclerosis; Osteopetrosis autosomal dominant type 1 and 2,

recessive 4, recessive 1, recessive 6; Osteoporosis with pseudoglioma; Oto-palato-digital syndrome, types I and II; Ovarian dysgenesis 1; Ovarioleukodystrophy; Pachyonychia congenita 4 and type 2; Paget disease of bone, familial; Pallister-Hall syndrome; Palmoplantar keratoderma, nonepidermolytic, focal or diffuse; Pancreatic agenesis and congenital heart disease; Papillon-Lef\xc3\xa8vre syndrome; Paragangliomas 3; Paramyotonia congenita of von Eulenburg; Parathyroid carcinoma; Parkinson disease 14, 15, 19 (juvenile-onset), 2, 20 (early-onset), 6, (autosomal recessive early-onset, and 9; Partial albinism; Partial hypoxanthine-guanine phosphoribosyltransferase deficiency; Patterned dystrophy of retinal pigment epithelium; PC-K6a; Pelizaeus-Merzbacher disease; Pendred syndrome; Peripheral demyelinating neuropathy, central dysmyelination; Hirschsprung disease; Permanent neonatal diabetes mellitus; Diabetes mellitus, permanent neonatal, with neurologic features; Neonatal insulin-dependent diabetes mellitus; Maturity-onset diabetes of the young, type 2; Peroxisome biogenesis disorder 14B, 2A, 4A, 5B, 6A, 7A, and 7B; Perrault syndrome 4; Perry syndrome; Persistent hyperinsulinemic hypoglycemia of infancy; familial hyperinsulinism; Phenotypes; Phenylketonuria; Pheochromocytoma; Hereditary Paraganglioma-Pheochromocytoma Syndromes; Paragangliomas 1; Carcinoid tumor of intestine; Cowden syndrome 3; Phosphoglycerate dehydrogenase deficiency; Phosphoglycerate kinase 1 deficiency; Photosensitive trichothiodystrophy; Phytanic acid storage disease; Pick disease; Pierson syndrome; Pigmentary retinal dystrophy; Pigmented nodular adrenocortical disease, primary, 1; Pilomatrixoma; Pitt-Hopkins syndrome; Pituitary dependent hypercortisolism; Pituitary hormone deficiency, combined 1, 2, 3, and 4; Plasminogen activator inhibitor type 1 deficiency; Plasminogen deficiency, type I; Platelet-type bleeding disorder 15 and 8; Poikiloderma, hereditary fibrosing, with tendon contractures, myopathy, and pulmonary fibrosis; Polycystic kidney disease 2, adult type, and infantile type; Polycystic lipomembranous osteodysplasia with sclerosing leukoencephalopathy; Polyglucosan body myopathy 1 with or without immunodeficiency; Polymicrogyria, asymmetric, bilateral frontoparietal; Polyneuropathy, hearing loss, ataxia, retinitis pigmentosa, and cataract; Pontocerebellar hypoplasia type 4; Popliteal pterygium syndrome; Porencephaly 2; Porokeratosis 8, disseminated superficial actinic type; Porphobilinogen synthase deficiency; Porphyria cutanea tarda; Posterior column ataxia with retinitis pigmentosa; Posterior polar cataract type 2; Prader-Willi-like syndrome; Premature ovarian failure 4, 5, 7, and 9; Primary autosomal recessive microcephaly 10, 2, 3, and 5; Primary ciliary dyskinesia 24; Primary dilated cardiomyopathy; Left ventricular noncompaction 6; 4, Left ventricular noncompaction 10; Paroxysmal atrial fibrillation; Primary hyperoxaluria,

type I, type, and type III; Primary hypertrophic osteoarthropathy, autosomal recessive 2; Primary hypomagnesemia; Primary open angle glaucoma juvenile onset 1; Primary pulmonary hypertension; Primrose syndrome; Progressive familial heart block type 1B; Progressive familial intrahepatic cholestasis 2 and 3; Progressive intrahepatic cholestasis; Progressive myoclonus epilepsy with ataxia; Progressive pseudorheumatoid dysplasia; Progressive sclerosing poliodystrophy; Prolidase deficiency; Proline dehydrogenase deficiency; Schizophrenia 4; Properdin deficiency, X-linked; Propionic academia; Proprotein convertase 1/3 deficiency; Prostate cancer, hereditary, 2; Protan defect; Proteinuria; Finnish congenital nephrotic syndrome; Proteus syndrome; Breast adenocarcinoma; Pseudoachondroplastic spondyloepiphyseal dysplasia syndrome; Pseudohypoaldosteronism type 1 autosomal dominant and recessive and type 2; Pseudohypoparathyroidism type 1A, Pseudopseudohypoparathyroidism; Pseudoneonatal adrenoleukodystrophy; Pseudoprimary hyperaldosteronism; Pseudoxanthoma elasticum; Generalized arterial calcification of infancy 2; Pseudoxanthoma elasticum-like disorder with multiple coagulation factor deficiency; Psoriasis susceptibility 2; PTEN hamartoma tumor syndrome; Pulmonary arterial hypertension related to hereditary hemorrhagic telangiectasia; Pulmonary Fibrosis And/Or Bone Marrow Failure, Telomere-Related, 1 and 3; Pulmonary hypertension, primary, 1, with hereditary hemorrhagic telangiectasia; Purine-nucleoside phosphorylase deficiency; Pyruvate carboxylase deficiency; Pyruvate dehydrogenase E1-alpha deficiency; Pyruvate kinase deficiency of red cells; Raine syndrome; Rasopathy; Recessive dystrophic epidermolysis bullosa; Nail disorder, nonsyndromic congenital, 8; Reifenstein syndrome; Renal adysplasia; Renal carnitine transport defect; Renal coloboma syndrome; Renal dysplasia; Renal dysplasia, retinal pigmentary dystrophy, cerebellar ataxia and skeletal dysplasia; Renal tubular acidosis, distal, autosomal recessive, with late-onset sensorineural hearing loss, or with hemolytic anemia; Renal tubular acidosis, proximal, with ocular abnormalities and mental retardation; Retinal cone dystrophy 3B; Retinitis pigmentosa; Retinitis pigmentosa 10, 11, 12, 14, 15, 17, and 19; Retinitis pigmentosa 2, 20, 25, 35, 36, 38, 39, 4, 40, 43, 45, 48, 66, 7, 70, 72; Retinoblastoma; Rett disorder; Rhabdoid tumor predisposition syndrome 2; Rhegmatogenous retinal detachment, autosomal dominant; Rhizomelic chondrodysplasia punctata type 2 and type 3; Roberts-SC phocomelia syndrome; Robinow Sorauf syndrome; Robinow syndrome, autosomal recessive, autosomal recessive, with brachy-syn-polydactyly; Rothmund-Thomson syndrome; Rapadilino syndrome; RRM2B-related mitochondrial disease; Rubinstein-Taybi syndrome; Salla disease; Sandhoff disease, adult and infantil types; Sarcoidosis, early-onset; Blau syndrome; Schindler disease, type 1; Schizencephaly;

Schizophrenia 15; Schneckenbecken dysplasia; Schwannomatosis 2; Schwartz Jampel syndrome type 1; Sclerocornea, autosomal recessive; Sclerosteosis; Secondary hypothyroidism; Segawa syndrome, autosomal recessive; Senior-Loken syndrome 4 and 5, ; Sensory ataxic neuropathy, dysarthria, and ophthalmoparesis; Sepiapterin reductase deficiency; SeSAME syndrome; Severe combined immunodeficiency due to ADA deficiency, with microcephaly, growth retardation, and sensitivity to ionizing radiation, atypical, autosomal recessive, T cell-negative, B cell-positive, NK cell-negative of NK-positive; Partial adenosine deaminase deficiency; Severe congenital neutropenia; Severe congenital neutropenia 3, autosomal recessive or dominant; Severe congenital neutropenia and 6, autosomal recessive; Severe myoclonic epilepsy in infancy; Generalized epilepsy with febrile seizures plus, types 1 and 2; Severe X-linked myotubular myopathy; Short QT syndrome 3; Short stature with nonspecific skeletal abnormalities; Short stature, auditory canal atresia, mandibular hypoplasia, skeletal abnormalities; Short stature, onychodysplasia, facial dysmorphism, and hypotrichosis; Primordial dwarfism; Short-rib thoracic dysplasia 11 or 3 with or without polydactyly; Sialidosis type I and II; Silver spastic paraplegia syndrome; Slowed nerve conduction velocity, autosomal dominant; Smith-Lemli-Opitz syndrome; Snyder Robinson syndrome; Somatotroph adenoma; Prolactinoma; familial, Pituitary adenoma predisposition; Sotos syndrome 1 or 2; Spastic ataxia 5, autosomal recessive, Charlevoix-Saguenay type, 1,10,or 11, autosomal recessive; Amyotrophic lateral sclerosis type 5; Spastic paraplegia 15, 2, 3, 35, 39, 4, autosomal dominant, 55, autosomal recessive, and 5A; Bile acid synthesis defect, congenital, 3; Spermatogenic failure 11, 3, and 8; Spherocytosis types 4 and 5; Spheroid body myopathy; Spinal muscular atrophy, lower extremity predominant 2, autosomal dominant; Spinal muscular atrophy, type II; Spinocerebellar ataxia 14, 21, 35, 40,and 6; Spinocerebellar ataxia autosomal recessive 1 and 16; Splenic hypoplasia; Spondylocarpotarsal synostosis syndrome; Spondylocheirodysplasia, Ehlers-Danlos syndrome-like, with immune dysregulation, Aggrecan type, with congenital joint dislocations, short limb-hand type, Sedaghatian type, with cone-rod dystrophy, and Kozlowski type; Parastremmatic dwarfism; Stargardt disease 1; Cone-rod dystrophy 3; Stickler syndrome type 1; Kniest dysplasia; Stickler syndrome, types 1(nonsyndromic ocular) and 4; Sting-associated vasculopathy, infantile-onset; Stormorken syndrome; Sturge-Weber syndrome, Capillary malformations, congenital, 1; Succinyl-CoA acetoacetate transferase deficiency; Sucrase-isomaltase deficiency; Sudden infant death syndrome; Sulfite oxidase deficiency, isolated; Supravalvar aortic stenosis; Surfactant metabolism dysfunction, pulmonary, 2 and 3; Symphalangism, proximal, 1b; Syndactyly Cenani Lenz type;

Syndactyly type 3; Syndromic X-linked mental retardation 16; Talipes equinovarus; Tangier disease; TARP syndrome; Tay-Sachs disease, B1 variant, Gm2-gangliosidosis (adult), Gm2-gangliosidosis (adult-onset); Temtamy syndrome; Tenorio Syndrome; Terminal osseous dysplasia; Testosterone 17-beta-dehydrogenase deficiency; Tetraamelia, autosomal recessive; Tetralogy of Fallot; Hypoplastic left heart syndrome 2; Truncus arteriosus; Malformation of the heart and great vessels; Ventricular septal defect 1; Thiel-Behnke corneal dystrophy; Thoracic aortic aneurysms and aortic dissections; Marfanoid habitus; Three M syndrome 2; Thrombocytopenia, platelet dysfunction, hemolysis, and imbalanced globin synthesis; Thrombocytopenia, X-linked; Thrombophilia, hereditary, due to protein C deficiency, autosomal dominant and recessive; Thyroid agenesis; Thyroid cancer, follicular; Thyroid hormone metabolism, abnormal; Thyroid hormone resistance, generalized, autosomal dominant; Thyrotoxic periodic paralysis and Thyrotoxic periodic paralysis 2; Thyrotropin-releasing hormone resistance, generalized; Timothy syndrome; TNF receptor-associated periodic fever syndrome (TRAPS); Tooth agenesis, selective, 3 and 4; Torsades de pointes; Townes-Brocks-branchiootorenal-like syndrome; Transient bullous dermolysis of the newborn; Treacher collins syndrome 1; Trichomegaly with mental retardation, dwarfism and pigmentary degeneration of retina; Trichorhinophalangeal dysplasia type I; Trichorhinophalangeal syndrome type 3; Trimethylaminuria; Tuberous sclerosis syndrome; Lymphangiomyomatosis; Tuberous sclerosis 1 and 2; Tyrosinase-negative oculocutaneous albinism; Tyrosinase-positive oculocutaneous albinism; Tyrosinemia type I; UDPglucose-4-epimerase deficiency; Ullrich congenital muscular dystrophy; Ulna and fibula absence of with severe limb deficiency; Upshaw-Schulman syndrome; Urocanate hydratase deficiency; Usher syndrome, types 1, 1B, 1D, 1G, 2A, 2C, and 2D; Retinitis pigmentosa 39; UV-sensitive syndrome; Van der Woude syndrome; Van Maldergem syndrome 2; Hennekam lymphangiectasia-lymphedema syndrome 2; Variegate porphyria; Ventriculomegaly with cystic kidney disease; Verheij syndrome; Very long chain acyl-CoA dehydrogenase deficiency; Vesicoureteral reflux 8; Visceral heterotaxy 5, autosomal; Visceral myopathy; Vitamin D-dependent rickets, types 1and 2; Vitelliform dystrophy ; von Willebrand disease type 2M and type 3; Waardenburg syndrome type 1, 4C, and 2E (with neurologic involvement); Klein-Waardenberg syndrome; Walker-Warburg congenital muscular dystrophy; Warburg micro syndrome 2 and 4; Warts, hypogammaglobulinemia, infections, and myelokathexis; Weaver syndrome; Weill-Marchesani syndrome 1 and 3; Weill-Marchesani-like syndrome; Weissenbacher-Zweymuller syndrome; Werdnig-Hoffmann disease; Charcot-Marie-Tooth disease; Werner syndrome; WFS1-Related Disorders;

Wiedemann-Steiner syndrome; Wilson disease; Wolfram-like syndrome, autosomal dominant; Worth disease; Van Buchem disease type 2; Xeroderma pigmentosum, complementation group b, group D, group E, and group G; X-linked agammaglobulinemia; X-linked hereditary motor and sensory neuropathy; X-linked ichthyosis with steryl-sulfatase deficiency; X-linked periventricular heterotopia; Oto-palato-digital syndrome, type I; X-linked severe combined immunodeficiency; Zimmermann-Laband syndrome and Zimmermann-Laband syndrome 2; and Zonular pulverulent cataract 3.

[00354] The instant disclosure provides lists of genes comprising pathogenic G to A or C to T mutations. Such pathogenic G to A or C to T mutations may be corrected using the methods and compositions provided herein, for example by mutating the A to a G, and/or the T to a C, thereby restoring gene function. Table 2 includes exemplary mutations that can be corrected using nucleobase editors provided herein. Table 2 includes the gene symbol, the associated phenotype, the mutation to be corrected and exemplary gRNA sequences which may be used to correct the mutations. The gRNA sequences provided in Table 2 are sequences that encode RNA that can direct Cas9, or any of the base editors provided herin, to a target site. For example, the gRNA sequences provided in Table 2 may be cloned into a gRNA expression vector, such as pFYF to encode a gRNA that targets Cas9, or any of the base editors provided herein, to a target site in order to correct a disease-related mutation. It should be appreciated, however, that additional mutations may be corrected to treat additional diseases associated with a G to A or C to T mutation. Furthermore, additional gRNAs may be designed based on the disclosure and the knowledge in the art, which would be appreciated by the skilled artisan.

Table 2 Target mutations that may be corrected using nucleobase editors, including exemplary gRNA sequences. gRNA sequences in the table correspond to SEQ ID NOs:740-5526 from top to bottom.

RS# (dbSNP)	Gene Symbol	Name	gRNAall	Phenotypes
80338761	SEPT9	NM_006640.4(SEPT9):c.262C>T (p.Arg88Trp)	CCGAGCCGGTGTCCYGGCGCACT	Hereditary neuralgic amyotrophy
80338762	SEPT9	NM_006640.4(SEPT9):c.278C>T (p.Ser93Phe)	CCCGGCGCACTGAGCTGTYCATT, CCGGCGCACTGAGCTGTYCATTG	Hereditary neuralgic amyotrophy
28934586	CYP11B1	NM_000497.3(CYP11B1):c.1343G>A (p.Arg448His)	GCATGCRCCAGTGCCTTGGGCGG	Deficiency of steroid 11-beta-monooxygenase
748979061	GRM6	NM_000843.3(GRM6):c.1462C>T (p.Gln488Ter)	TGCCTRGTACCCGCCACTGCTGG	Congenital stationary night blindness, type 1B
786205118	CHKB	NM_005198.4(CHKB):c.677+1G>A	AACCTCAGRTGAGGGCAGGCAGG	Muscular dystrophy, congenital, megaconial type
121965029	IDUA	NM_000203.4(IDUA):c.266G>A (p.Arg89Gln)	GGTCCRGACCCACTGGCTGCTGG	Mucopolysaccharidosis, MPS-I-H/S, Hurler syndrome
104893659	PAX8	NM_013953.3(PAX8):c.170G>A (p.Cys57Tyr)	GGCTRCGTCAGCAAGATCCTTGG	Thyroid agenesis
104894062	CYP11B1	NM_000497.3(CYP11B1):c.1121G>A (p.Arg374Gln)	CTTGCRGTGGGTGCTGGCTGAGG	Deficiency of steroid 11-beta-monooxygenase
104894231	HRAS	NM_005343.2(HRAS):c.436G>A (p.Ala146Thr)	GACCTCGRCCAAGACCCGGCAGG	Costello syndrome
104894335	AQP2	NM_000486.5(AQP2):c.523G>A (p.Gly175Arg)	CTTRGGGTAGGTCATGGCCATGG	
104894341	AQP2	NM_000486.5(AQP2):c.568G>A (p.Ala190Thr)	CTGRCTCCAGCTGTCGTCACTGG	
104894604	NAGS	NM_153006.2(NAGS):c.971G>A (p.Trp324Ter)	AGTRGGTGAGCACAAAAGAACGG	Hyperammonemia, type III
104894832	GLA	NM_000169.2(GLA):c.982G>A (p.Gly328Arg)	GCAARGGTACCAGCTTAGACAGG	Fabry disease
104894842	GLA	NM_000169.2(GLA):c.1020G>A (p.Trp340Ter)	GTGTGRGAACGACCTCTCTCAGG	Fabry disease
794726859	SLC6A1	NM_003042.3(SLC6A1):c.131G>A (p.Arg44Gln)	CCCGACCRGGACACGTGGAAGGG, CCCCGACCRGGACACGTGGAAGG	MYOCLONIC-ATONIC EPILEPSY
794727219	HADHA	NM_000182.4(HADHA):c.2146+1G>A	GGAGRTTGGTCTCGCAGGTTGGG, GGGAGRTTGGTCTCGCAGGTTGG	Mitochondrial trifunctional protein deficiency, Long-chain 3-hydroxyacyl-CoA dehydrogenase deficiency

672601330	ACY1	NM_000666.2(ACY1):c.360-1G>A	TCARGTACCTGGAAGCTGTGAGG	Aminoacylase 1 deficiency
794727995	COPA	NM_001098398.1(COPA):c.721G>A (p.Glu241Lys)	TGGRAGGTTGATACCTGCCGGGG, ATGGRAGGTTGATACCTGCCGGG, CATGGRAGGTTGATACCTGCCGG	
794728029	ACTA2	NM_001613.2(ACTA2):c.809G>A (p.Gly270Glu)	TTTCCAGRGATGGAGTCTGCTGG	Thoracic aortic aneurysms and aortic dissections
397514667	CD27	NM_001242.4(CD27):c.158G>A (p.Cys53Tyr)	GGACTRTGACCAGCATAGAAAGG	Lymphoproliferative syndrome 2
397514668	GDF5	NM_000557.4(GDF5):c.1139G>A (p.Arg380Gln)	GCGAAAACRGCGGGCCCCACTGG	Brachydactyly type A2
794729274	TTN	NM_001256850.1(TTN):c.49637G>A (p.Trp16546Ter)	AGTGAGCTRGACTCCTCCTTTGG	not provided
377461670	MYH7	NM_000257.3(MYH7):c.5029C>T (p.Arg1677Cys)	GTTGTTGCRCCGCTCCACGATGG	Cardiomyopathy
794729383	TTN	NM_001256850.1(TTN):c.71708G>A (p.Trp23903Ter)	GTTAAATRGGGAAAGGTGGATGG	not provided
121909129	KRT86	NM_002284.3(KRT86):c.1237G>A (p.Glu413Lys)	GGAGGGCRAGGAGCAGAGGTGGG, TGGAGGGCRAGGAGCAGAGGTGG	Beaded hair, not provided
397516208	MYH7	NM_000257.3(MYH7):c.4276G>A (p.Glu1426Lys)	TGAGATCRAGGACTTGATGGTGG	Cardiomyopathy, not specified
397516211	MYH7	NM_000257.3(MYH7):c.4348G>A (p.Asp1450Asn)	ACTTCRACAAGGTGGGCCCTGGG, AACTTCRACAAGGTGGGCCCTGG	Cardiomyopathy, not specified
386834035	POMGNT1	NM_017739.3(POMGNT1):c.652+1G>A	AAAGGAGRTGCCGGCATCAGAGG	Muscle eye brain disease, Congenital muscular dystrophy-dystroglycanopathy with mental retardation, type B3
752034900	ACO2	NM_001098.2(ACO2):c.1981G>A (p.Gly661Arg)	GTGATCRGAGACGAGAACTACGG	Optic atrophy 9
57419521	KRT81	NM_002281.3(KRT81):c.1237G>A (p.Glu413Lys)	GGAGGGCRAGGAGCAGAGGTGGG, TGGAGGGCRAGGAGCAGAGGTGG	Beaded hair, not provided
774122562	RDH5	NM_002905.3(RDH5):c.285G>A (p.Trp95Ter)	GTGRGTGGAGATGCACGTTAAGG	Pigmentary retinal dystrophy
121908508	SCO2	NM_001169109.1(SCO2):c.	TCCTRGCTTTTGTCAAGGCAGGG,	Cardioencephalomyopathy, fatal infantile, due to cytochrome

		107G>A (p.Trp36Ter)	GTCCTRGCTTTTGTCAAGGCAGG	c oxidase deficiency
199473684	GLA	NM_000169.2(GLA):c.639 +919G>A	CTARAGTGTAAGTTTCATGAGGG, ACTARAGTGTAAGTTTCATGAGG	Fabry disease, Fabry disease, cardiac variant
121912591	NAGS	NM_153006.2(NAGS):c.83 5G>A (p.Ala279Thr)	GGTGACCRCGTCGCTGGCCAAGG	Hyperammonemia, type III
121912826	COL4A3	NM_000091.4(COL4A3):c. 3044G>A (p.Gly1015Glu)	CCAGRAAGCATGGGGAACATGG G, ACCAGRAAGCATGGGGAACATG G	Benign familial hematuria
730880914	MYH7	NM_000257.3(MYH7):c.50 30G>A (p.Arg1677His)	GCGGCRCAACAACCTGCTGCAGG	Cardiomyopathy
121912932	COL5A1	NM_000093.4(COL5A1):c. 4466G>A (p.Gly1489Glu)	ATCGRGCTCATCGGTCCTCCGGG, GATCGRGCTCATCGGTCCTCCGG	Ehlers-Danlos syndrome, classic type
199474822		m.7444G>A	AAAATCTARACAAAAAAGGAAG G	Leber optic atrophy, Aminoglycoside-induced deafness, Deafness, nonsyndromic sensorineural, mitochondrial
398123050	RTEL1	NM_016434.3(RTEL1):c.21 41+5G>A	GGTGCRTGCAGTCCGGTGGCAGG	Dyskeratosis congenita, autosomal recessive, 5
115556836	PTCH1	NM_000264.3(PTCH1):c.2 183C>T (p.Thr728Met)	ACTTCRTACAGGGGGGCTCGAGG	Holoprosencephaly 7, not specified, not provided
201540674	RTEL1	RTEL1:c.2402G>A (p.p.Arg801His)	TCCAGCRCTGCCAAGCCTGCTGG	Idiopathic fibrosing alveolitis, chronic form, Dyskeratosis congenita, autosomal recessive, 5, PULMONARY FIBROSIS AND/OR BONE MARROW FAILURE, TELOMERE-RELATED, 3
62638208	GRM6	NM_000843.3(GRM6):c.15 65G>A (p.Cys522Tyr)	TGCCCTRCGGGCCGGGGGAGCGG	Congenital stationary night blindness, type 1B, not provided
62638625	GRM6	NM_000843.3(GRM6):c.23 41G>A (p.Glu781Lys)	TTCAACRAGGCCAAGCCCATCGG	Congenital stationary night blindness, type 1B, not provided
796053008	SCN1A	NM_001165963.1(SCN1A): c.4285G>A (p.Ala1429Thr)	AATAGRCCACATTCAAAGGATGG	not provided
121917756	HRAS	NM_005343.2(HRAS):c.18 7G>A (p.Glu63Lys)	GGAGRAGTACAGCGCCATGCGG G, AGGAGRAGTACAGCGCCATGCG G	Myopathy, congenital, with excess of muscle spindles
121917775	VIM	NM_003380.3(VIM):c.451 G>A (p.Glu151Lys)	CCTCTACRAGGAGGAGATGCGGG , ACCTCTACRAGGAGGAGATGCGG	Cataract, nuclear diffuse nonprogressive, not provided
121917995	SCN1A	NM_006920.4(SCN1A):c.4 874G>A (p.Arg1625Gln)	TCCRAGTGATCCGTCTTGCTAGG	Generalized epilepsy with febrile seizures plus, type 2, Epileptic encephalopathy Lennox-Gastaut type, not provided
121918805	SCN1A	NM_006920.4(SCN1A):c.4	ATGGGCRTAAATTTGTTTGCTGG	Generalized epilepsy with febrile seizures plus, type 1, not

		063G>A (p.Val1355Ile)		provided
768431507	TTN	NM_001256850.1(TTN):c.49243C>T (p.Arg16415Ter)	CATTTCRGAACACTGAGCCAAGG	not provided
368138001	NPHP3	NM_153240.4(NPHP3):c.3373C>T (p.Arg1125Ter)	AACTCRCTCCCTCATTTCTAAGG	Adolescent nephronophthisis, Renal-hepatic-pancreatic dysplasia, not provided
730880440	GLA	NM_000169.2(GLA):c.1019G>A (p.Trp340Ter)	GTGTRGGAACGACCTCTCTCAGG	not provided
730880450	GLA	NM_000169.2(GLA):c.713G>A (p.Ser238Asn)	GAAAARTATAAAGAGTATCTTGG	not provided
267606745	COL4A3	NM_000091.4(COL4A3):c.3499G>A (p.Gly1167Arg)	GCCRGAGAAAAGGGAGAAACGGG, AGCCRGAGAAAAGGGAGAAACGG	Alport syndrome, autosomal dominant
267606961	POMGNT1	NM_001243766.1(POMGNT1):c.1425G>A (p.Trp475Ter)	GGGATTGRGACATGTGGATGCGG	
267607132	TOP1	NM_003286.2(TOP1):c.1748G= (p.Gly583=)	CATGGAGGRCTTGACAGCCAAGG	
398122933	CD27	NM_001242.4(CD27):c.24G>A (p.Trp8Ter)	CCTGGTGRCTGTGCGTTCTGGGG, CCCTGGTGRCTGTGCGTTCTGGG	Lymphoproliferative syndrome 2
398122944	CRYGC	NM_020989.3(CRYGC):c.471G>A (p.Trp157Ter)	CTGRGGGGCCATGGATGCTAAGG	Cataract, coppock-like
398123019	RTEL1	NM_032957.4(RTEL1):c.823G>A (p.Glu275Lys)	CTTTGACRAAGCTCACAACGTGG	Dyskeratosis congenita, autosomal recessive, 5
281865424	TYRP1	NM_000550.2(TYRP1):c.1067G>A (p.Arg356Gln)	AGTTTCCRAAACACAGTGGAAGG	Oculocutaneous albinism type 3
111033564	PRSS1	NM_002769.4(PRSS1):c.235G>A (p.Glu79Lys)	ACAACATCRAAGTCCTGGAGGGG	Hereditary pancreatitis
121434424	GDF1	NM_001492.5(GDF1):c.485G>A (p.Gly162Asp)	GGGCGRCTGGGAGCTGAGCGTGG	Tetralogy of Fallot, not provided
74315389	GDF5	NM_000557.4(GDF5):c.1471G>A (p.Glu491Lys)	GCAGTATRAGGACATGGTCGTGG	Symphalangism, proximal, 1b
74315511	SCO2	NM_005138.2(SCO2):c.418G>A (p.Glu140Lys)	CCCAGACRAGCTGGAGAAGCTGG	Myopia 6, Cardioencephalomyopathy, fatal infantile, due to cytochrome c oxidase deficiency, not provided
137853207	DDC	NM_001082971.1(DDC):c.304G>A (p.Gly102Ser)	CATCRGCTTCTCCTGGGTGAGGG, GCATCRGCTTCTCCTGGGTGAGG	Deficiency of aromatic-L-amino-acid decarboxylase
786201031	SPECC1L	NM_015330.4(SPECC1L):c.3247G>A (p.Gly1083Ser)	TGTCRGCATCAAATCCACACTGG	Opitz G/BBB syndrome
387906858	KCNJ13	NM_002242.4(KCNJ13):c.4	AAAAGCTCRATTTTTTGGCCGGG	Leber congenital amaurosis 16

		96C>T (p.Arg166Ter)		
387907183	ALG11	NM_001004127.2(ALG11):c.1192G>A (p.Glu398Lys)	TGGAACRAGCATTTTGGGATTGG	Congenital disorder of glycosylation type 1P
72646831	TTN	NM_001267550.2(TTN):c.57331C>T (p.Arg19111Ter)	CCATGCTGGAGGGGTGATCYGAA	Primary dilated cardiomyopathy, Dilated cardiomyopathy 1G, not provided
750586158	RAD50	NM_005732.3(RAD50):c.3598C>T (p.Arg1200Ter)	CCTTGGATATGCGAGGAYGATGC	Hereditary cancer-predisposing syndrome
34516117	KCNQ1	NM_000218.2(KCNQ1):c.1799C>T (p.Thr600Met)	CCTTTGTCCCCGCAGGTGAYGCA	Long QT syndrome, Congenital long QT syndrome, Cardiac arrhythmia
786205700	RTEL1	NM_032957.4(RTEL1):c.1523C>T (p.Pro508Leu)	CCAGCGGCACGCTGGCCCYGGTG	PULMONARY FIBROSIS AND/OR BONE MARROW FAILURE, TELOMERE-RELATED, 3
786205701	RTEL1	NM_032957.4(RTEL1):c.2149C>T (p.Gln717Ter)	CCAGGGCTGTGAACYAGGCCATC	PULMONARY FIBROSIS AND/OR BONE MARROW FAILURE, TELOMERE-RELATED, 3
104893940	ARG1	NM_000045.3(ARG1):c.871C>T (p.Arg291Ter)	CCAGAAGAAGTAACTYGAACAGT	Arginase deficiency
104894069	CYP11B1	NM_000497.3(CYP11B1):c.124C>T (p.Pro42Ser)	CCCTTTGAAGCCATGYCCCGGCG, CCTTTGAAGCCATGYCCCGGCGT	Congenital adrenal hyperplasia, Deficiency of steroid 11-beta-monooxygenase
145100473	SCO2	NM_001169109.1(SCO2):c.341G>A (p.Arg114His)	CCGGAAGTCAGCCTTGCAGYGAG	Myopia 6
104894333	AQP2	NM_000486.5(AQP2):c.374C>T (p.Thr125Met)	CCCCAGCTCAGCAACAGCAYGAC, CCCAGCTCAGCAACAGCAYGACG, CCAGCTCAGCAACAGCAYGACGG	
794726710	SCN1A	NM_001165963.1(SCN1A):c.3637C>T (p.Arg1213Ter)	CCTGAGAAGGACGTGTTTCYGAA	Severe myoclonic epilepsy in infancy, not provided
794726752	SCN1A	NM_001165963.1(SCN1A):c.4573C>T (p.Arg1525Ter)	CCGCAAAAGCCTATACCTYGACC	Severe myoclonic epilepsy in infancy, not provided
794726759	SCN1A	NM_001165963.1(SCN1A):c.4933C>T (p.Arg1645Ter)	CCGTCTTGCTAGGATTGGCYGAA	Severe myoclonic epilepsy in infancy
104894837	GLA	NM_000169.2(GLA):c.436C>T (p.Pro146Ser)	CCTGCGCAGGCTTCYCTGGGAGT	Fabry disease
200970763	PIEZO1	NM_001142864.3(PIEZO1):c.2344G>A (p.Gly782Ser)	CCGCTCAGCCACCAGGCYCCACT	Xerocytosis
794726988	RAD51D	NM_002878.3(RAD51D):c.955C>T (p.Gln319Ter)	CCTGGGGGACCTCAGAGYAGAGT	Breast-ovarian cancer, familial 4
794728015	SOX18	NM_018419.2(SOX18):c.481C>T (p.Gln161Ter)	CCGGCCGCGCCGCAAGAAGYAGG,	Hypotrichosis-lymphedema-telangiectasia syndrome

			CCGCGCCGCAAGAAGYAGGCGC G	
794728540	KCNQ1	NM_000218.2(KCNQ1):c.1801C>T (p.Gln601Ter)	CCCCGCAGGTGACGYAGCTGGAC	Cardiac arrhythmia
28940869	POMGNT1	NM_017739.3(POMGNT1):c.1324C>T (p.Arg442Cys)	CCCAGCACTACTGTACYGTGTGG, CCAGCACTACTGTACYGTGTGGA	Congenital muscular dystrophy
794729279	TTN	NM_001256850.1(TTN):c.58702C>T (p.Arg19568Ter)	CCCCTAAAGTCACTTGGYGAAAA , CCCTAAAGTCACTTGGYGAAAAG , CCTAAAGTCACTTGGYGAAAAGT	not provided
794729305	TTN	NM_001256850.1(TTN):c.96304C>T (p.Arg32102Ter)	CCGTGTAGGACAGGCCYGAGAA A	not provided
794729384	TTN	NM_001256850.1(TTN):c.81193C>T (p.Arg27065Ter)	CCATTCAGAGCTTAYGAGGGACA	not provided
397517326	CDH23	NM_022124.5(CDH23):c.3628C>T (p.Gln1210Ter)	CCCCCGTGTTCACAYAGCAGCAG	Usher syndrome, type 1D
147708782	MYH2	NM_017534.5(MYH2):c.706G>A (p.Ala236Thr)	CCTCACGGTCTTGGYGTTGCCAA	Inclusion body myopathy 3
121908189	MFRP	NM_031433.3(MFRP):c.523C>T (p.Gln175Ter)	CCACTGCGTGTGGCATATCYAGG	Nanophthalmos 2
121908350	CDH23	NM_022124.5(CDH23):c.3880C>T (p.Gln1294Ter)	CCAGGCCCCGCCCTTCAACYAGG , CCCCGCCCTTCAACYAGGGCTTC	Usher syndrome, type 1D
121908490	SGCE	NM_003919.2(SGCE):c.304C>T (p.Arg102Ter)	CCGACCTGGATGGCTTYGATATA	Myoclonic dystonia
121909106	FOXC2	NM_005251.2(FOXC2):c.374C>T (p.Ser125Leu)	CCGCCACAACCTCTYGCTCAACG	Distichiasis-lymphedema syndrome
121909117	SOX10	NM_006941.3(SOX10):c.470C>T (p.Ala157Val)	CCCCTTCATCGAGGAGGYTGAGC , CCCTTCATCGAGGAGGYTGAGCG , CCTTCATCGAGGAGGYTGAGCGG	Waardenburg syndrome type 4C
121909501	TDGF1	NM_001174136.1(TDGF1):c.326C>T (p.Pro109Leu)	CCCCATGACACCTGGCTGCYCAA, CCCATGACACCTGGCTGCYCAAG , CCATGACACCTGGCTGCYCAAGA	Forebrain defects
121909511	CHRNE	NM_000080.3(CHRNE):c.865C>T (p.Leu289Phe)	CCAGACCGTCTTCTTGTTCYTCA, CCGTCTTCTTGTTCYTCATTGCC	Myasthenia, familial infantile, 1

121909512	CHRNE	NM_000080.3(CHRNE):c.422C>T (p.Pro141Leu)	CCGTGACGTGGCTGCCTCYGGCC	MYASTHENIC SYNDROME, CONGENITAL, 4B, FAST-CHANNEL
121909595	CRYGD	NM_006891.3(CRYGD):c.43C>T (p.Arg15Cys)	CCGGGGCTTCCAGGGCYGCCACT	Cataract 4
121912420	SERPIND1	NM_000185.3(SERPIND1):c.1385C>T (p.Pro462Leu)	CCACGGTGGGGTTCATGCYGCTG	Heparin cofactor II deficiency
111033571	TRIM32	NM_012210.3(TRIM32):c.388C>T (p.Pro130Ser)	CCGGGAGGCAGACCATCAGYCTC	Bardet-Biedl syndrome, Bardet-Biedl syndrome 11
121912700	ACY1	NM_000666.2(ACY1):c.589C>T (p.Arg197Trp)	CCTCCCCAGGGGTGYGGGTTACC	Aminoacylase 1 deficiency
121912978	CYP11B2	NM_000498.3(CYP11B2):c.554C>T (p.Thr185Ile)	CCCGGGGGAGCCTGAYCCTGGAC, CCGGGGGAGCCTGAYCCTGGACG	Corticosterone methyloxidase type 2 deficiency
730882050	ALG14	NM_144988.3(ALG14):c.194C>T (p.Pro65Leu)	CCAATGCCTACTCACYTAGACAT	Myasthenic syndrome, congenital, without tubular aggregates
761807131	TTN	NM_001256850.1(TTN):c.46513+1G>A	CCATGTCCAAACTTAYGCTTTGG	not provided
730882144	MFRP	NM_031433.3(MFRP):c.1549C>T (p.Arg517Trp)	CCCTGCTACCAGCATTTCYGGAG, CCTGCTACCAGCATTTCYGGAGG	Microphthalmia, isolated 5
61752068	RS1	NM_000330.3(RS1):c.305G>A (p.Arg102Gln)	CCTTGACTGTTGAGCYGGGCCTT	Juvenile retinoschisis, not provided
119456962	NPHP3	NM_153240.4(NPHP3):c.1729C>T (p.Arg577Ter)	CCTCCTTGATTATTAAAYGACTA, CCTTGATTATTAAAYGACTAACT	Adolescent nephronophthisis, Renal-hepatic-pancreatic dysplasia, Meckel syndrome type 7, not provided
74315310	FCGR1A	NM_000566.3(FCGR1A):c.274C>T (p.Arg92Ter)	CCAGAGAGGTCTCTCAGGGYGAA	
41469351	CCR5	NM_000579.3(CCR5):c.-229C>T	CCCGTAAATAAACCTTYAGACCA, CCGTAAATAAACCTTYAGACCAG	
62638185	RDH5	NM_002905.3(RDH5):c.218C>T (p.Ser73Phe)	CCTGCAGCGGGTGGCCTYCTCCC	
796053004	SCN1A	NM_001165963.1(SCN1A):c.3985C>T (p.Arg1329Ter)	CCTCTAAGAGCCTTATCTYGATT	Severe myoclonic epilepsy in infancy, not provided
796053103	SCN1A	NM_001165963.1(SCN1A):c.5710C>T (p.Gln1904Ter)	CCTTCCAAGGTCTCCTATYAGCC, CCAAGGTCTCCTATYAGCCAATC	not provided
121917886	P2RY12	NM_022788.4(P2RY12):c.793C>T (p.Arg265Trp)	CCCTGAGCCAAACCYGGGATGTC	Platelet-type bleeding disorder 8
121917903	ERCC6	NM_000124.3(ERCC6):c.229C>T (p.Arg77Ter)	CCCTGCTGCACATCGACYGACAT, CCTGCTGCACATCGACYGACATC	UV-sensitive syndrome
121918450	ITGB3	NM_000212.2(ITGB3):c.22	CCTCATCACCATCCACGACYGAA	Glanzmann thrombasthenia

		48C>T (p.Arg750Ter)		
727503949	GLA	NM_000169.2(GLA):c.658 C>T (p.Arg220Ter)	CCCAATTATACAGAAATCYGACA , CCAATTATACAGAAATCYGACAG	Fabry disease
202102042	NPPA	NM_006172.3(NPPA):c.44 9G>A (p.Arg150Gln)	CCCCAGTTCCTCTTACCYGGAAG, CCCAGTTCCTCTTACCYGGAAGC, CCAGTTCCTCTTACCYGGAAGCT	Atrial standstill 2
202003805	PRSS1	NM_002769.4(PRSS1):c.47 C>T (p.Ala16Val)	CCACTCCAGTTGCTGYCCCCTTT	Hereditary pancreatitis
587781698	ATM	NM_000051.3(ATM):c.899 8C>T (p.Gln3000Ter)	CCTTAGTGATATTGACYAGAGTT	Hereditary cancer-predisposing syndrome
148231754	TTN	NM_001256850.1(TTN):c.8 3086+5G>A	CCCTTCCCATGACAAATAYGTAC, CCTTCCCATGACAAATAYGTACC	not provided
587781756	RAD51D	NM_002878.3(RAD51D):c. 451C>T (p.Gln151Ter)	CCGCCTCCTCCAGCTGCTTYAGG, CCTCCTCCAGCTGCTTYAGGCTA	Hereditary cancer-predisposing syndrome
137853289	ABCA12	NM_173076.2(ABCA12):c. 6610C>T (p.Arg2204Ter)	CCATGTTTTTTTCCTTGYGACTC	Autosomal recessive congenital ichthyosis 4B
116840809	RPL35A	NM_000996.2(RPL35A):c. 304C>T (p.Arg102Ter)	CCATTGGACACAGAATCYGAGTG	Diamond-Blackfan anemia 5
587782695	RAD51D	NM_002878.3(RAD51D):c. 547C>T (p.Gln183Ter)	CCAGATGCTGGATGTGCTGYAGG	Hereditary cancer-predisposing syndrome
119462978	KIRREL3	NM_032531.3(KIRREL3):c .118C>T (p.Arg40Trp)	CCAAGGACAAGTTTYGGAGAATG	Mental retardation, autosomal dominant 4
398123812	SGCE	NM_003919.2(SGCE):c.70 9C>T (p.Arg237Ter)	CCCGTTTTCTTCTTGTTTAYGAG, CCGTTTTCTTCTTGTTTAYGAGA	not provided
397518485	TRAF3IP 2	NM_147686.3(TRAF3IP2): c.1580C>T (p.Thr527Ile)	CCCACCTGGCTTCAGAACAYTCA, CCACCTGGCTTCAGAACAYTCAT, CCTGGCTTCAGAACAYTCATGTC	Candidiasis, familial, 8
794726839	SCN1A	NM_001165963.1(SCN1A): c.4985C>T (p.Ala1662Val)	CCGCACGCTGCTCTTTGYTTTGA	Severe myoclonic epilepsy in infancy
60035576	KRT10	NM_000421.3(KRT10):c.13 00C>T (p.Gln434Ter)	CCAGAATACTGAATACCAAYAAC	Bullous ichthyosiform erythroderma, not provided
779874042	TTN	NM_001256850.1(TTN):c.7 7716G>T (p.Glu25906Ter)	CCCACACCAGCTGCATTTTHAGC, CCACACCAGCTGCATTTTHAGCA	not provided
121434219	ATM	NM_000051.3(ATM):c.913 9C>T (p.Arg3047Ter)	CCCCAAAAATCTCAGCYGACTTT, CCCAAAAATCTCAGCYGACTTTT, CCAAAAATCTCAGCYGACTTTTC	Ataxia-telangiectasia syndrome
121434306	LPAR6	NM_005767.5(LPAR6):c.4 63C>T (p.Gln155Ter)	CCCGCCGTTTTTGTTYAGTCTAC, CCGCCGTTTTTGTTYAGTCTACC	Hypotrichosis 8

121434405	RPL5	NM_000969.3(RPL5):c.67C>T (p.Arg23Ter)	CCAAGTGAAATTTAGAAGAYGAC	Aase syndrome
121434410	PRKRA	NM_003690.4(PRKRA):c.665C>T (p.Pro222Leu)	CCTTGAGGAATTCTCYTGGTGAA	Dystonia 16
121434534	CYP19A1	NM_031226.2(CYP19A1):c.1303C>T (p.Arg435Cys)	CCATTTGGCTTTGGGCCCYGTGG	Aromatase deficiency
121434571	ERCC5	NM_000123.3(ERCC5):c.2375C>T (p.Ala792Val)	CCCATGGAAGCAGAGGYGCAGTG, CCATGGAAGCAGAGGYGCAGTGC	Xeroderma pigmentosum, group G
137852293	PHKA2	NM_000292.2(PHKA2):c.3341C>T (p.Thr1114Ile)	CCCTTTGGTAGATGAYCCCGCAT, CCTTTGGTAGATGAYCCCGCATG	Glycogen storage disease IXa2
74315359	PINK1	NM_032409.2(PINK1):c.938C>T (p.Thr313Met)	CCTGGGCCATGGCCGGAYGCTGT	Parkinson disease 6, autosomal recessive early-onset
372635387	CHRNE	NM_000080.3(CHRNE):c.37G>A (p.Gly13Arg)	CCGTACCGAGAAGCCYCAAGAGG	MYASTHENIC SYNDROME, CONGENITAL, 4B, FAST-CHANNEL
74315520	SOX10	NM_006941.3(SOX10):c.1129C>T (p.Gln377Ter)	CCAGCCATCCACCTCAYAGATCG	Waardenburg syndrome type 4C, Waardenburg syndrome type 2E, with neurologic involvement
74315521	SOX10	NM_006941.3(SOX10):c.748C>T (p.Gln250Ter)	CCCCGAAGACAGAGCTGYAGTCG, CCCGAAGACAGAGCTGYAGTCGG, CCGAAGACAGAGCTGYAGTCGGG	Peripheral demyelinating neuropathy, central dysmyelination, Waardenburg syndrome, and Hirschsprung disease
137852824	PCSK1	NM_000439.4(PCSK1):c.920C>T (p.Ser307Leu)	CCATCTTCGTCTGGGCTTYGGGA	Proprotein convertase 1/3 deficiency
137852973	BSCL2	NM_001122955.3(BSCL2):c.461C>T (p.Ser154Leu)	CCCTGTTGCCAATGTCTYGCTGA, CCTGTTGCCAATGTCTYGCTGAC	Silver spastic paraplegia syndrome
137852974	BSCL2	NM_001122955.3(BSCL2):c.1015C>T (p.Arg339Ter)	CCTCCACAGGTTAACATCYGAAA, CCACAGGTTAACATCYGAAAAAG	Congenital generalized lipodystrophy type 2
137853211	DDC	NM_001082971.1(DDC):c.272C>T (p.Ala91Val)	CCCGGCCATGCTTGYGGACATGC	Deficiency of aromatic-L-amino-acid decarboxylase
387906570	APOA1	NM_000039.1(APOA1):c.67C>T (p.Gln23Ter)	CCAGGCTCGGCATTTCTGGYAGC	Tangier disease
387906669	ACP5	NM_001111035.1(ACP5):c.667C>T (p.Gln223Ter)	CCCACTGCCTGGTCAAGYAGCTA, CCACTGCCTGGTCAAGYAGCTAC	Spondyloenchondrodysplasia with immune dysregulation
80051519	CYP19A1	NM_031226.2(CYP19A1):c.1094G>A (p.Arg365Gln)	CCACGACAGGCTGGTACYGCATG	Aromatase deficiency

387907268	PRCD	NM_001077620.2(PRCD):c.64C>T (p.Arg22Ter)	CCGCCGATTTGCCAACYGAGTCC	Retinitis pigmentosa 36
387907279	FAN1	NM_014967.4(FAN1):c.2245C>T (p.Arg749Ter)	CCGCCTTTCACTGTATCAGYGAG, CCTTTCACTGTATCAGYGAGCCG	Interstitial nephritis, karyomegalic
543267101	AARS2	NM_020745.3(AARS2):c.2893G>A (p.Gly965Arg)	CCAGGTCAGTAGTGCTTCYGGTG	Leukoencephalopathy, progressive, with ovarian failure
587777592	AARS2	NM_020745.3(AARS2):c.1213G>A (p.Glu405Lys)	CCAGGAAGGCTGCCTYGTCCTCT	Leukoencephalopathy, progressive, with ovarian failure
587777125	AASS	NM_005763.3(AASS):c.194G>A (p.Arg65Gln)	CCTTATCATGAATGGCCYGCCGA	Hyperlysinemia
121434578	ABAT	NM_000663.4(ABAT):c.659G>A (p.Arg220Lys)	GGGARGACCATGGGTAAGGAGGG, TGGGARGACCATGGGTAAGGAGG, CCATGGGARGACCATGGGTAAGG	Gamma-aminobutyric acid transaminase deficiency
724159990	ABAT	NM_000663.4(ABAT):c.631C>T (p.Leu211Phe)	CCCCGACTACAGCATCYTCTCCT, CCCGACTACAGCATCYTCTCCTT, CCGACTACAGCATCYTCTCCTTC	Gamma-aminobutyric acid transaminase deficiency
137854495	ABCA1	NM_005502.3(ABCA1):c.2810C>T (p.Ala937Val)	CCTGGGCCACAATGGAGYGGGGA	Tangier disease
28940270	ABCA12	NM_173076.2(ABCA12):c.4541G>A (p.Arg1514His)	ATGTTCTCRCCGAAGTATATGGG	Autosomal recessive congenital ichthyosis 4A
121909181	ABCA3	NM_001089.2(ABCA3):c.3426G>A (p.Trp1142Ter)	CTGRCTCTCTGCTCTGCTGTGGG, TCTGRCTCTCTGCTCTGCTGTGG	Surfactant metabolism dysfunction, pulmonary, 3
61750061	ABCA4	NM_000350.2(ABCA4):c.3106G>A (p.Glu1036Lys)	GTCCCAGRAGGAGGCCCAGCTGG	Stargardt disease 1, not provided
61751399	ABCA4	NM_000350.2(ABCA4):c.3364G>A (p.Glu1122Lys)	GGACRAGGCCGACCTCCTTGGGG, TGGACRAGGCCGACCTCCTTGGG, ATGGACRAGGCCGACCTCCTTGG	Stargardt disease 1, not provided
1800553	ABCA4	NM_000350.2(ABCA4):c.5882G>A (p.Gly1961Glu)	TGTCGRAGTTCGCCCTGGAGAGG	Stargardt disease 1, Cone-rod dystrophy 3, not provided
794727531	ABCA4	NM_000350.2(ABCA4):c.4429C>T (p.Gln1477Ter)	CCCAGCTGTTCCAGAAGYAGAAA, CCAGCTGTTCCAGAAGYAGAAAT	Stargardt disease 1
794727903	ABCA4	NM_000350.2(ABCA4):c.880C>T (p.Gln294Ter)	CCATCGGCCGAGTATGYAGGACT	Stargardt disease 1, Cone-rod dystrophy 3
61748550	ABCA4	NM_000350.2(ABCA4):c.1222C>T (p.Arg408Ter)	CCTGATTCACCTGCAGCAYGAAG	Retinitis pigmentosa 19, Stargardt disease 1, Cone-rod dystrophy 3, Age-related macular degeneration 2, not provided

61750130	ABCA4	NM_000350.2(ABCA4):c.4139C>T (p.Pro1380Leu)	CCCACAGATCGTGCTCCYGGCTA, CCACAGATCGTGCTCCYGGCTAC	Stargardt disease 1, not provided
28938473	ABCA4	NM_000350.2(ABCA4):c.5908C>T (p.Leu1970Phe)	CCTAGTGCTTTGGCYTCCTGGGA	Stargardt disease, not provided
72549401	ABCB11	NM_003742.2(ABCB11):c.1723C>T (p.Arg575Ter)	GGGATTTCRGATGAGGGCTCTGG	Progressive familial intrahepatic cholestasis 2
72552778	ABCB4	NM_018849.2(ABCB4):c.959C>T (p.Ser320Phe)	CCTTCTGGTATGGATYCACTCTA	Cholecystitis, Progressive familial intrahepatic cholestasis 3, Cholestasis, intrahepatic, of pregnancy 3
121918440	ABCB4	NM_018849.2(ABCB4):c.2869C>T (p.Arg957Ter)	CCTATGCCGGTTGTTTTYGATTT	Progressive familial intrahepatic cholestasis 3, Cholestasis, intrahepatic, of pregnancy 3
121918442	ABCB4	NM_018849.2(ABCB4):c.3502C>T (p.Pro1168Ser)	CCTTTCATCGAGACGTTAYCCCA	Cholecystitis
72558200	ABCC2	NM_000392.4(ABCC2):c.3449G>A (p.Arg1150His)	GGCRTCTGGACTCTGTCACCAGG	Dubin-Johnson syndrome
63749823	ABCC6	NM_001171.5(ABCC6):c.3961G>A (p.Gly1321Ser)	GAGRGTGGGATCTGGATCGACGG	Pseudoxanthoma elasticum
28939701	ABCC6	NM_001171.5(ABCC6):c.3412C>T (p.Arg1138Trp)	CCAGGGCAGCACAGTGGTCYGGG	Pseudoxanthoma elasticum
72653744	ABCC6	NM_001171.5(ABCC6):c.3490C>T (p.Arg1164Ter)	CCAGAGGATCAGTTTCCCGYGAC	Pseudoxanthoma elasticum
63750459	ABCC6	NM_001171.5(ABCC6):c.3389C>T (p.Thr1130Met)	CCCACATGGCTGAGAYGTTCCAG, CCACATGGCTGAGAYGTTCCAGG	Pseudoxanthoma elasticum
63750759	ABCC6	NM_001171.5(ABCC6):c.3940C>T (p.Arg1314Trp)	CCTGGCCAGTGGGCTGCTGYGGC, CCAGTGGGCTGCTGYGGCTCCAG	Pseudoxanthoma elasticum, Generalized arterial calcification of infancy 2
193922402	ABCC8	NM_000352.4(ABCC8):c.4306C>T (p.Arg1436Ter)	CCTCTTCAGCGGCACCATCYGGT	Persistent hyperinsulinemic hypoglycemia of infancy, familial hyperinsulinism
387906805	ABCC9	NM_005691.3(ABCC9):c.4640C>T (p.Thr1547Ile)	CCACTTTGGTGATGAYCAACAAG	Atrial fibrillation, familial, 12
387907228	ABCC9	NM_005691.3(ABCC9):c.3346C>T (p.Arg1116Cys)	CCTTGGAATCTCTAACTYGCTCA	Hypertrichotic osteochondrodysplasia
11146842	ABCD1	NM_000033.3(ABCD1):c.1850G>A (p.Arg617His)	TGGCCCRCATGTTCTACCACAGG	Adrenoleukodystrophy
150346282	ABCD1	NM_000033.3(ABCD1):c.1825G>A (p.Glu609Lys)	GGTGGCRAGAAGCAGAGAATCGG	Adrenoleukodystrophy
128624213	ABCD1	NM_000033.3(ABCD1):c.871G>A (p.Glu291Lys)	TCGRAGGAGATCGCCTTCTATGG	Adrenoleukodystrophy
128624218	ABCD1	NM_000033.3(ABCD1):c.7	GTTCRGGGAGCTGGTGGCAGAGG	Adrenoleukodystrophy

		96G>A (p.Gly266Arg)		
398123102	ABCD1	NM_000033.3(ABCD1):c.1553G>A (p.Arg518Gln)	TCCRGATCCTGGGTGGGCTCTGG	Adrenoleukodystrophy
398123107	ABCD1	NM_000033.3(ABCD1):c.1802G>A (p.Trp601Ter)	GACTRGAAGGACGTCCTGTCGGG, TGACTRGAAGGACGTCCTGTCGG	Adrenoleukodystrophy
398123105	ABCD1	NM_000033.3(ABCD1):c.1679C>T (p.Pro560Leu)	CCAGGTGATCTACCYGGACTCAG	Adrenoleukodystrophy
398123106	ABCD1	NM_000033.3(ABCD1):c.1771C>T (p.Arg591Trp)	CCTGCACCACATCCTGCAGYGGG	Adrenoleukodystrophy
587777603	ABHD12	NM_015600.4(ABHD12):c.477G>A (p.Trp159Ter)	CCAAGGCATCCTCATAYCACATC	Polyneuropathy, hearing loss, ataxia, retinitis pigmentosa, and cataract
387906517	ABL1	NM_007313.2(ABL1):c.763G>A (p.Glu255Lys)	CAAGTGGRAGATGGAACGCACGG	
121913459	ABL1	NM_007313.2(ABL1):c.1001C>T (p.Thr334Ile)	CCCCCGTTCTATATCATCAYTGA, CCCCGTTCTATATCATCAYTGAG, CCCGTTCTATATCATCAYTGAGT, CCGTTCTATATCATCAYTGAGTT	
368949613	ACAD9	NM_014049.4(ACAD9):c.1249C>T (p.Arg417Cys)	CCGTACGAGCGCATACTGYGTGA	Acyl-CoA dehydrogenase family, member 9, deficiency of
121434278	ACADM	NM_000016.5(ACADM):c.583G>A (p.Gly195Arg)	CCAACRGAGGAAAAGCTAATTGG	Medium-chain acyl-coenzyme A dehydrogenase deficiency, not provided
1799958	ACADS	NM_000017.3(ACADS):c.625G>A (p.Gly209Ser)	TCAGRGCATCAGTGCCTTCCTGG	Deficiency of butyryl-CoA dehydrogenase, not specified, not provided
387906951	ACADS	NM_000017.3(ACADS):c.323G>A (p.Gly108Asp)	AGCCGTGRCTGCGCCTCCACCGG	Deficiency of butyryl-CoA dehydrogenase
28940872	ACADS	NM_000017.3(ACADS):c.1147C>T (p.Arg383Cys)	CCGGCAGAGCGGCACTACYGCGA	Deficiency of butyryl-CoA dehydrogenase, not provided
28940874	ACADS	NM_000017.3(ACADS):c.575C>T (p.Ala192Val)	CCAATGCCTGGGAGGYTTCGGCT	Deficiency of butyryl-CoA dehydrogenase
28941773	ACADS	NM_000017.3(ACADS):c.1058C>T (p.Ser353Leu)	CCAAGCTGGCCGCCTYGGAGGCC	Deficiency of butyryl-CoA dehydrogenase, not provided
121908006	ACADS	NM_000017.3(ACADS):c.973C>T (p.Arg325Trp)	CCTGGCCCTGGAGAGTGCCYGGC, CCCTGGAGAGTGCCYGGCTGCTG	Deficiency of butyryl-CoA dehydrogenase, not provided
61732144	ACADS	NM_000017.3(ACADS):c.319C>T (p.Arg107Cys)	CCATGGAGGAGATCAGCYGTGGC	Deficiency of butyryl-CoA dehydrogenase, not provided
147442301	ACADS	NM_000017.3(ACADS):c.164C>T (p.Pro55Leu)	CCGAGAAGGAGTTGTTTCYCATT	Deficiency of butyryl-CoA dehydrogenase

188094280	ACADSB	NM_001609.3(ACADSB):c.1159G>A (p.Glu387Lys)	ATCRAGTGGATGGGGGGAGTAGG	Deficiency of 2-methylbutyryl-CoA dehydrogenase
137852649	ACADSB	NM_001609.3(ACADSB):c.763C>T (p.Leu255Phe)	CCTGAAAACAAATTGGGGYTCAG	Deficiency of 2-methylbutyryl-CoA dehydrogenase
533055438	ACADVL	NM_000018.3(ACADVL):c.1316G>A (p.Gly439Asp)	GGGGGRTATGGGCTTCATGAAGG	not provided
2309689	ACADVL	NM_000018.3(ACADVL):c.1322G>A (p.Gly441Asp)	TATGGRCTTCATGAAGGTACAGG	Very long chain acyl-CoA dehydrogenase deficiency, not provided
766742117	ACADVL	NM_000018.3(ACADVL):c.1375C>T (p.Arg459Trp)	CCGAGATCTTCGCATCTTCYGGA	not provided
118204014	ACADVL	NM_000018.3(ACADVL):c.1837C>T (p.Arg613Trp)	CCCCAGGCTGCAGCTYGGATCCG, CCCAGGCTGCAGCTYGGATCCGA	Very long chain acyl-CoA dehydrogenase deficiency, not provided
121913568	ACAN	NM_013227.3(ACAN):c.7141G>A (p.Asp2381Asn)	GAACRACAGGACCATCGAAGGG, TGAACRACAGGACCATCGAAGGG, CTGAACRACAGGACCATCGAAGG	Spondyloepimetaphyseal dysplasia, Aggrecan type
267606625	ACAN	NM_013227.3(ACAN):c.7249G>A (p.Val2417Met)	AGGACTGTRTGGTGATGATCTGG	Osteochondritis dissecans
121912703	ACE	NM_000789.3(ACE):c.3683C>T (p.Pro1228Leu)	CCGCAGTACAACTGGACGCYGAA	Angiotensin i-converting enzyme, benign serum increase
118204093	ACOX1	NM_004035.6(ACOX1):c.442C>T (p.Arg148Ter)	CCTTAGGAACTCACCTTYGAGGC	Pseudoneonatal adrenoleukodystrophy
757905943	ACSF3	NM_174917.4(ACSF3):c.348G>A (p.Trp116Ter)	GTCATGRATGAGTGGCGGTGTGG	not provided
138680796	ACSF3	NM_174917.4(ACSF3):c.1411C>T (p.Arg471Trp)	CCAGTACTGGATCCGAGGCYGGA	Combined malonic and methylmalonic aciduria
140986055	ACSF3	NM_174917.4(ACSF3):c.728C>T (p.Pro243Leu)	CCTCCACGTGCTCCYGCTGCACC	Combined malonic and methylmalonic aciduria
367543051	ACTA1	NM_001100.3(ACTA1):c.727G>A (p.Glu243Lys)	CTACRAGCTGCCAGACGGGCAGG, AGAGCTACRAGCTGCCAGACGGG	Congenital myopathy with fiber type disproportion
267606627	ACTA1	NM_001100.3(ACTA1):c.223C>T (p.His75Tyr)	CCTGAAGTACCCTATCGAGYACG	Nemaline myopathy 3
794728021	ACTA2	NM_001613.2(ACTA2):c.116G>A (p.Arg39His)	GACRTCCCAGACATCAGGTGAGG, TGTGGGACRTCCCAGACATCAGG	Thoracic aortic aneurysms and aortic dissections
387906592	ACTA2	NM_001613.2(ACTA2):c.5	ATCATGCRTCTGGATCTGGCTGG	Aortic aneurysm, familial thoracic 6, Thoracic aortic

		36G>A (p.Arg179His)		aneurysms and aortic dissections, Multisystemic smooth muscle dysfunction syndrome, Moyamoya disease 5
112901682	ACTA2	NM_001141945.1(ACTA2):c.115C>T (p.Arg39Cys)	CCCATCCATTGTGGGABGTCCCA, CCATCCATTGTGGGABGTCCCAG	Aortic aneurysm, familial thoracic 6, Thoracic aortic aneurysms and aortic dissections
397515470	ACTB	NM_001101.3(ACTB):c.349G>A (p.Glu117Lys)	CAACCGCRAGAAGATGACCCAGG	Baraitser-Winter syndrome 1
587779770	ACTB	NM_001101.3(ACTB):c.220G>A (p.Gly74Ser)	GCACRGCATCGTCACCAACTGGG, AGCACRGCATCGTCACCAACTGG	Baraitser-Winter syndrome 1
587779769	ACTB	NM_001101.3(ACTB):c.209C>T (p.Pro70Leu)	CCTCACCCTGAAGTACCYCATCG	Baraitser-Winter syndrome 1
587779774	ACTB	NM_001101.3(ACTB):c.359C>T (p.Thr120Ile)	CCAACCGCGAGAAGATGAYCCAG, CCGCGAGAAGATGAYCCAGGTGA	Baraitser-Winter syndrome 1
587779775	ACTB	NM_001101.3(ACTB):c.446C>T (p.Thr149Ile)	CCTCTGGCCGTACCAYTGGCATC	Baraitser-Winter syndrome 1
104894546	ACTG1	NM_001614.3(ACTG1):c.791C>T (p.Pro264Leu)	CCGGAGGCGCTGTTCCAGCYTTC	Deafness, autosomal dominant 20
281875325	ACTG1	NM_001614.3(ACTG1):c.359C>T (p.Thr120Ile)	CCAACAGAGAGAAGATGAYTCAG	Baraitser-Winter Syndrome 2, not provided
11549190	ACTG1	NM_001614.3(ACTG1):c.404C>T (p.Ala135Val)	CCCCGGCCATGTACGTGGYCATC, CCCGGCCATGTACGTGGYCATCC, CCGGCCATGTACGTGGYCATCCA	Baraitser-Winter Syndrome 2, not provided
78001248	ACTG2	NM_001615.3(ACTG2):c.532C>T (p.Arg178Cys)	CCTGCCCCATGCCATCATGYGCC, CCCCATGCCATCATGYGCCTGGA, CCCATGCCATCATGYGCCTGGAC	Visceral myopathy
587777385	ACTG2	NM_001615.3(ACTG2):c.118C>T (p.Arg40Cys)	CCATTGTGGGCCGCCCTYGCCAC	Visceral myopathy
387907345	ACTN1	NM_001130004.1(ACTN1):c.313G>A (p.Val105Ile)	GGCRTCAAACTGGTGTCCATCGG	Platelet-type bleeding disorder 15
794728966	ACTN2	NM_001103.3(ACTN2):c.2527-1G>A	CCARCCATACATCCTGGCGGAGG, TTCCCARCCATACATCCTGGCGG	Cardiomyopathy
727502886	ACTN2	NM_001103.3(ACTN2):c.355G>A (p.Ala119Thr)	GGCRCTGAAGGTGAGAGGTGTGG, CCATTGGCRCTGAAGGTGAGAGG	Dilated cardiomyopathy 1AA, Primary familial hypertrophic cardiomyopathy, Familial hypertrophic cardiomyopathy 23
121434437	ACVR2B	NM_001106.3(ACVR2B):c.119G>A (p.Arg40His)	CTGGAGCRCACCAACCAGAGCGG	Heterotaxy, visceral, 4, autosomal
28936687	ACVRL1	NM_000020.2(ACVRL1):c.	AAGRCCGCTATGGCGAAGTGTGG	Pulmonary arterial hypertension related to hereditary

		632G>A (p.Gly211Asp)		hemorrhagic telangiectasia
28936688	ACVRL1	NM_000020.2(ACVRL1):c.1031G>A (p.Cys344Tyr)	CAGTRTTGCATCGCCGACCTGGG, GCAGTRTTGCATCGCCGACCTGG	Hereditary hemorrhagic telangiectasia type 2, Pulmonary arterial hypertension related to hereditary hemorrhagic telangiectasia
28936401	ACVRL1	NM_000020.2(ACVRL1):c.1120C>T (p.Arg374Trp)	CCCGAGAGTGGGCACCAAGYGGT, CCGAGAGTGGGCACCAAGYGGTA	Hereditary hemorrhagic telangiectasia type 2, Pulmonary arterial hypertension related to hereditary hemorrhagic telangiectasia
121909288	ACVRL1	NM_000020.2(ACVRL1):c.1450C>T (p.Arg484Trp)	CCCGACTCACCGCGCTGYGGATC, CCGACTCACCGCGCTGYGGATCA	Hereditary hemorrhagic telangiectasia type 2, Pulmonary arterial hypertension related to hereditary hemorrhagic telangiectasia
121908716	ADA	NM_000022.2(ADA):c.632G>A (p.Arg211His)	TCACCRTACTGTCCACGCCGGGG, TTCACCRTACTGTCCACGCCGGG, ATTCACCRTACTGTCCACGCCGG	Severe combined immunodeficiency due to ADA deficiency
121908723	ADA	NM_000022.2(ADA):c.646G>A (p.Gly216Arg)	CCACGCCRGGGAGGTGGGCTCGG	Severe combined immunodeficiency due to ADA deficiency
121908715	ADA	NM_000022.2(ADA):c.986C>T (p.Ala329Val)	CCTTTCCAGAACATCAATGYGGC, CCAGAACATCAATGYGGCCAAAT	Severe combined immunodeficiency due to ADA deficiency
121908735	ADA	NM_000022.2(ADA):c.466C>T (p.Arg156Cys)	CCATCCTGTGCTGCATGYGCCAC	Severe combined immunodeficiency due to ADA deficiency
121908736	ADA	NM_000022.2(ADA):c.226C>T (p.Arg76Trp)	CCCTTCCCAGGGGCTGCYGGGAG, CCTTCCCAGGGGCTGCYGGGAGG	Severe combined immunodeficiency due to ADA deficiency, Partial adenosine deaminase deficiency
114025668	ADA	NM_000022.2(ADA):c.643G>A (p.Ala215Thr)	CCGAGCCCACCTCCCCGGYGTGG	Severe combined immunodeficiency due to ADA deficiency, Partial adenosine deaminase deficiency
121434358	ADAMTS10	NM_030957.3(ADAMTS10):c.73G>A (p.Ala25Thr)	ACGCACRCCTTCCGGTCTCAAGG	Weill-Marchesani syndrome 1
121434357	ADAMTS10	NM_030957.3(ADAMTS10):c.709C>T (p.Arg237Ter)	CCTGAAGCGATCGGTCAGCYGAG	Weill-Marchesani syndrome 1
786205077	ADAMTS13	NM_139025.4(ADAMTS13):c.414+1G>A	GCCTGAGRTAGGCATGGAGCTGG	Upshaw-Schulman syndrome
281875305	ADAMTS13	NM_139025.4(ADAMTS13):c.1523G>A (p.Cys508Tyr)	GGTRTATGCCAAGTGGCCCCCGG	Upshaw-Schulman syndrome, not provided
121908471	ADAMTS13	NM_139025.4(ADAMTS13):c.1193G>A (p.Arg398His)	TGCTCCCRCTCCTGCGGAGGAGG	Upshaw-Schulman syndrome
121908474	ADAMTS13	NM_139025.4(ADAMTS13):c.3638G>A (p.Cys1213Tyr)	ACTRTGCAGTGGCCATTGGGCGG, CAGACTRTGCAGTGGCCATTGGG, GCAGACTRTGCAGTGGCCATTGG	Upshaw-Schulman syndrome
786205078	ADAMTS	NM_139025.4(ADAMTS13	ACARGGGGCAGAACTGCTTCGGG	Upshaw-Schulman syndrome

	13	):c.331-1G>A	, CACARGGGGCAGAACTGCTTCGG	
11575933	ADAMTS 13	NM_139025.4(ADAMTS13):c.1423C>T (p.Pro475Ser)	CCACTGGGGTGCTGCTGTAYCAC	Upshaw-Schulman syndrome
121908469	ADAMTS 13	NM_139025.4(ADAMTS13):c.304C>T (p.Arg102Cys)	CCAGGAGGACACAGAGYGCTAT G	Upshaw-Schulman syndrome
121908478	ADAMTS 13	NM_139025.4(ADAMTS13):c.749C>T (p.Ala250Val)	CCCAGCGGACACGTGATGGYTTC , CCAGCGGACACGTGATGGYTTCG	Upshaw-Schulman syndrome
267606638	ADAMTS 17	NM_139057.2(ADAMTS17):c.760C>T (p.Gln254Ter)	CCACGGGGCCGAGGCCGCCYAG A	Weill-Marchesani-like syndrome
137853147	ADAMTS 2	NM_014244.4(ADAMTS2) :c.2384G>A (p.Trp795Ter)	GGAGTRGGAGTACAGAGACGAG G	Ehlers-Danlos syndrome type 7, autosomal recessive
137853146	ADAMTS 2	NM_014244.4(ADAMTS2) :c.673C>T (p.Gln225Ter)	CCCTCCTCTCGGGGGGCCAYAGG , CCTCCTCTCGGGGGGCCAYAGGC , CCTCTCGGGGGGCCAYAGGCCCT	Ehlers-Danlos syndrome type 7, autosomal recessive
387907064	ADAMTS L2	NM_001145320.1(ADAMT SL2):c.215G>A (p.Arg72Gln)	AGCRGCACTGCCTGCAGCAGAGG	Acromicric dysplasia
113994121	ADAMTS L2	NM_001145320.1(ADAMT SL2):c.440C>T (p.Pro147Leu)	CCACATCTCCAGCAAACYGTGTG	Acromicric dysplasia
387907065	ADAMTS L2	NM_001145320.1(ADAMT SL2):c.661C>T (p.Arg221Cys)	CCACGTGACGGGCAACTATYGCA	Acromicric dysplasia
368482584	ADAMTS L4	NM_019032.5(ADAMTSL 4):c.2008C>T (p.Arg670Ter)	CCAGCTGCGTACTGGAAAYGAGT	Ectopia lentis, isolated autosomal recessive
119468004	ADCK3	NM_020247.4(ADCK3):c.1 651G>A (p.Glu551Lys)	CTACRAGGTCAAGGTGAGCAGGG , GCTACRAGGTCAAGGTGAGCAGG	Coenzyme Q10 deficiency, primary, 4
119468009	ADCK3	NM_020247.4(ADCK3):c.1 645G>A (p.Gly549Ser)	CCTCACCRGCTACGAGGTCAAGG	Coenzyme Q10 deficiency, primary, 4
119468005	ADCK3	NM_020247.4(ADCK3):c.6 37C>T (p.Arg213Trp)	CCTGTGACGAGGATTGGCYGGCT	Coenzyme Q10 deficiency, primary, 4
398122981	ADCK4	NM_024876.3(ADCK4):c.1 027C>T (p.Arg343Trp)	CCTAAGCCAGGACCTGYGGAACC	Nephrotic syndrome, type 9

587777497	ADCY1	NM_021116.2(ADCY1):c.3112C>T (p.Arg1038Ter)	CCCCTACCACTTTGTGTGCYGAG, CCCTACCACTTTGTGTGCYGAGG, CCTACCACTTTGTGTGCYGAGGC	Deafness, autosomal recessive 44
757156390	ADCY5	NM_183357.2(ADCY5):c.1425C>G (p.Ile475Met)	AGCCCTCRATGTCAGCAAACAGG	Multiple congenital anomalies
796065306	ADCY5	NM_183357.2(ADCY5):c.2176G>A (p.Ala726Thr)	TTGACRCCAGGAGCATTGATAGG	Dyskinesia, familial, with facial myokymia
587783657	ADGRG1	NM_005682.6(ADGRG1):c.1970G>A (p.Trp657Ter)	TCATCTRGTACTGGTCCATGCGG	Polymicrogyria, bilateral frontoparietal
587783660	ADGRG1	NM_005682.6(ADGRG1):c.620+1G>A	GCCARTAAGTTTGGCACCTGGGG, AGCCARTAAGTTTGGCACCTGGG, CAGCCARTAAGTTTGGCACCTGG	Polymicrogyria, bilateral frontoparietal
121908464	ADGRG1	NM_005682.6(ADGRG1):c.1693C>T (p.Arg565Trp)	CCCGCAGGTGCTGGATCYGGGAC, CCGCAGGTGCTGGATCYGGGACT	Polymicrogyria, bilateral frontoparietal
587783658	ADGRG1	NM_005682.6(ADGRG1):c.265C>T (p.His89Tyr)	CCCCAGGGGCCTCTACYACTTCT, CCCAGGGGCCTCTACYACTTCTG, CCAGGGGCCTCTACYACTTCTGC	Polymicrogyria, bilateral frontoparietal
121909762	ADGRV1	NM_032119.3(ADGRV1):c.6901C>T (p.Gln2301Ter)	CCCCTGGGGAAACCATTYAAACC, CCCTGGGGAAACCATTYAAACCT, CCTGGGGAAACCATTYAAACCTT	Usher syndrome, type 2C
28941471	ADSL	NM_000026.2(ADSL):c.569G>A (p.Arg190Gln)	TCCRAGATGACCTGCGCTTCCGG	Adenylosuccinate lyase deficiency, not provided
202064195	ADSL	NM_000026.2(ADSL):c.953C>T (p.Pro318Leu)	CCCTTGTCATGGACCYGCTACAG, CCTTGTCATGGACCYGCTACAGA	not provided
750614500	ADSL	NM_000026.2(ADSL):c.568C>T (p.Arg190Ter)	CCAGAACTTGAAGCGTGTCYGAG	not provided
756210458	ADSL	NM_000026.2(ADSL):c.421C>T (p.Arg141Trp)	CCAGAGTGATCTCTYGGCTTGCC	not provided
786205680	AFF4	NM_014423.3(AFF4):c.772C>T (p.Arg258Trp)	CCCACTGCCTATGTGYGGCCCAT, CCACTGCCTATGTGYGGCCCATG	CHOPS SYNDROME
727502823	AFG3L2	NM_006796.2(AFG3L2):c.1875G>A (p.Met625Ile)	AGGATRTGTATGACTTTAGGTGG, GATAGGATRTGTATGACTTTAGG	Spastic ataxia 5, autosomal recessive
121434412	AGPS	NM_003659.3(AGPS):c.926C>T (p.Thr309Ile)	CCCTGGAGTTCAGTAYTGTAGGA, CCTGGAGTTCAGTAYTGTAGGAG	Rhizomelic chondrodysplasia punctata type 3
74315283	AGT	NM_000029.3(AGT):c.1124G>A (p.Arg375Gln)	TCTCCCCRGTAGGAGCCTCCCGG	Renal dysplasia

121908522	AGXT	NM_000030.2(AGXT):c.245G>A (p.Gly82Glu)	CTCGGRACACTGTGCCCTGGAGG, TGGCTCGGRACACTGTGCCCTGG	Primary hyperoxaluria, type I
121908523	AGXT	NM_000030.2(AGXT):c.121G>A (p.Gly41Arg)	GCCRGGGGGCTGCAGATGATCGG	Primary hyperoxaluria, type I
121908528	AGXT	NM_000030.2(AGXT):c.738G>A (p.Trp246Ter)	AAGTGRCTGGCCAACTTCTGGGG, CAAGTGRCTGGCCAACTTCTGGG, TCAAGTGRCTGGCCAACTTCTGG	Primary hyperoxaluria, type I
121908530	AGXT	NM_000030.2(AGXT):c.466G>A (p.Gly156Arg)	ACCCACRGGGAGTCGTCCACCGG	Primary hyperoxaluria, type I
180177161	AGXT	NM_000030.2(AGXT):c.1079G>A (p.Arg360Gln)	GTGCTGCRGATCGGCCTGCTGGG, GGTGCTGCRGATCGGCCTGCTGG	Primary hyperoxaluria, type I
180177162	AGXT	NM_000030.2(AGXT):c.107G>A (p.Arg36His)	CTCRCATCATGGCAGCCGGGGGG , CCTCRCATCATGGCAGCCGGGGG , TCCTCRCATCATGGCAGCCGGGG, CTCCTCRCATCATGGCAGCCGGG, CCTCCTCRCATCATGGCAGCCGG	Primary hyperoxaluria, type I
180177163	AGXT	NM_000030.2(AGXT):c.1102G>A (p.Ala368Thr)	CAATRCCACCCGCGAGAATGTGG	Primary hyperoxaluria, type I
180177170	AGXT	NM_000030.2(AGXT):c.125G>A (p.Gly42Glu)	CCGGGGRGCTGCAGATGATCGGG , GCCGGGGRGCTGCAGATGATCGG	Primary hyperoxaluria, type I
180177177	AGXT	NM_000030.2(AGXT):c.166-1G>A	GCARATCATGGACGAGATCAAGG	Primary hyperoxaluria, type I
180177196	AGXT	NM_000030.2(AGXT):c.308G>A (p.Gly103Glu)	TTGRGGCCAATGGCATTTGGGGG, GTTGRGGCCAATGGCATTTGGGG, GGTTGRGGCCAATGGCATTTGGG, TGGTTGRGGCCAATGGCATTTGG	Primary hyperoxaluria, type I
180177198	AGXT	NM_000030.2(AGXT):c.323G>A (p.Trp108Ter)	CATTTRGGGGCAGCGAGCCGTGG	Primary hyperoxaluria, type I
180177231	AGXT	NM_000030.2(AGXT):c.518G>A (p.Cys173Tyr)	AACTCTRCCACAGGTGAGCCTGG	Primary hyperoxaluria, type I
180177235	AGXT	NM_000030.2(AGXT):c.533G>A (p.Cys178Tyr)	GTACAAGTRCCTGCTCCTGGTGG	Primary hyperoxaluria, type I
180177236	AGXT	NM_000030.2(AGXT):c.547G>A (p.Asp183Asn)	GTGRATTCGGTGGCATCCCTGGG, GGTGRATTCGGTGGCATCCCTGG	Primary hyperoxaluria, type I
61729604	AGXT	NM_000030.2(AGXT):c.866G>A (p.Arg289His)	CAGCTGGCRCCAGCACCGCGAGG	Primary hyperoxaluria, type I, not provided

180177210	AGXT	NM_000030.2(AGXT):c.364C>T (p.Arg122Ter)	CCTGCACCCAGGAGCCYGAGTGC	Primary hyperoxaluria, type I
180177279	AGXT	NM_000030.2(AGXT):c.844C>T (p.Gln282Ter)	CCTGGCCCTCATTGCGGAAYAGG, CCCTCATTGCGGAAYAGGTGCAT	Primary hyperoxaluria, type I
180177296	AGXT	NM_000030.2(AGXT):c.922C>T (p.Gln308Ter)	CCTGCAGGCACTGGGGCTGYAGC	Primary hyperoxaluria, type I
104894325	AICDA	NM_020661.2(AICDA):c.203G>A (p.Trp68Ter)	TCGGACTRGGACCTAGACCCTGG	Immunodeficiency with hyper IgM type 2
104894324	AICDA	NM_020661.2(AICDA):c.70C>T (p.Arg24Trp)	CCGCTGGGCTAAGGGTYGGCGTG	Immunodeficiency with hyper IgM type 2
104894190	AIP	NM_003977.3(AIP):c.911G>A (p.Arg304Gln)	GAGCCRAGAGCTGCAGGCCCTGG	Pituitary dependent hypercortisolism
104894194	AIP	NM_003977.3(AIP):c.40C>T (p.Gln14Ter)	CCGGGAGGACGGGATCYAAAAAC	Somatotroph adenoma, Prolactinoma, familial, Pituitary adenoma predisposition
267606541	AIP	NM_003977.3(AIP):c.241C>T (p.Arg81Ter)	CCATCGTGTGCACCATGYGAGAA	Somatotroph adenoma
62637014	AIPL1	NM_014336.4(AIPL1):c.834G>A (p.Trp278Ter)	GTGRAATGAGGCCGAGGCCAAGG	Leber congenital amaurosis 4, not provided
61757484	AIPL1	NM_014336.4(AIPL1):c.1126C>T (p.Pro376Ser)	GACGRGGGTGGCTCTGTGGCTGG, TGGGGACGRGGGTGGCTCTGTGG	Leber congenital amaurosis 4, not specified, not provided
142326926	AIPL1	NM_014336.4(AIPL1):c.784G>A (p.Gly262Ser)	CCCCTGCAGCCCCGCGCACYTGG, CCCTGCAGCCCCGCGCACYTGGG, CCTGCAGCCCCGCGCACYTGGGT	Leber congenital amaurosis 4, not provided
137853204	AK1	NM_000476.2(AK1):c.118G>A (p.Gly40Arg)	TCTCCACCRGGGACCTCCTGCGG	Adenylate kinase deficiency, hemolytic anemia due to
137853205	AK1	NM_000476.2(AK1):c.190G>A (p.Gly64Arg)	GAAGRGGCAGCTGGTTCCACTGG	Adenylate kinase deficiency, hemolytic anemia due to
104894101	AK1	NM_000476.2(AK1):c.382C>T (p.Arg128Trp)	CCCTGAGACCATGACCCAGYGGC, CCTGAGACCATGACCCAGYGGCT	Adenylate kinase deficiency, hemolytic anemia due to
121918343	AKR1D1	NM_005989.3(AKR1D1):c.316C>T (p.Leu106Phe)	CCCAACCCTGGAGAGGACAYTCA, CCAACCCTGGAGAGGACAYTCAG, CCCTGGAGAGGACAYTCAGGGTC	Bile acid synthesis defect, congenital, 2
121434592	AKT1	NM_005163.2(AKT1):c.49	TAGGGRAGTACATCAAGACCTGG	Proteus syndrome, Carcinoma of colon, Breast

		G>A (p.Glu17Lys)		adenocarcinoma, Neoplasm of ovary
121434593	AKT2	NM_001626.5(AKT2):c.821G>A (p.Arg274His)	TACCRCGACATCAAGGTTAGTGG	Diabetes mellitus type 2
387906659	AKT2	NM_001626.5(AKT2):c.49G>A (p.Glu17Lys)	CAGGTRAATACATCAAGACCTGG	
397514606	AKT3	NM_181690.2(AKT3):c.49G>A (p.Glu17Lys)	TAGGARAATATATAAAAAACTGG	Megalencephaly-polymicrogyria-polydactyly-hydrocephalus syndrome 2
121912981	ALAD	NM_000031.5(ALAD):c.823G>A (p.Val275Met)	CTCGCCRTGTACCACGTCTCTGG	Porphobilinogen synthase deficiency
121912983	ALAD	NM_000031.5(ALAD):c.820G>A (p.Ala274Thr)	CTCRCCGTGTACCACGTCTCTGG	Porphobilinogen synthase deficiency
749066913	ALAD	NM_000031.5(ALAD):c.165-11C>T	GGTRTGGGTAGAGGGGTTGAAGG	Porphobilinogen synthase deficiency
121912982	ALAD	NM_000031.5(ALAD):c.718C>T (p.Arg240Trp)	CCCTCATCCCTTAGGACYGGGAT, CCTCATCCCTTAGGACYGGGATG	Porphobilinogen synthase deficiency
137852302	ALAS2	NM_000032.4(ALAS2):c.871G>A (p.Gly291Ser)	ATCCAARGTATCCGTAACAGTGG	Hereditary sideroblastic anemia
386834230	ALDH1A3	NM_000693.3(ALDH1A3):c.211G>A (p.Val71Met)	CGACRTGGACAAGGCTGTGGAGG, GCCCGACRTGGACAAGGCTGTGG	not provided
397514652	ALDH1A3	NM_000693.3(ALDH1A3):c.265C>T (p.Arg89Cys)	CCAGAGGGGCTCGCCATGGYGCC	Microphthalmia, isolated 8
28939378	ALG1	NM_019109.4(ALG1):c.773C>T (p.Ser258Leu)	CCCAGTCACGGAGCGGTYGGCCT, CCAGTCACGGAGCGGTYGGCCTT	Congenital disorder of glycosylation type 1K, not provided
121907933	ALG12	NM_024105.3(ALG12):c.301G>A (p.Gly101Arg)	GTTAGARGAGTGCTTGGACTCGG	Congenital disorder of glycosylation type 1G
121907931	ALG12	NM_024105.3(ALG12):c.200C>T (p.Thr67Met)	CCCGGAGTCGTCCCCAGGAYGTT, CCGGAGTCGTCCCCAGGAYGTTC	Congenital disorder of glycosylation type 1G
367570129	ALG14	NM_144988.3(ALG14):c.310C>T (p.Arg104Ter)	AATTCRGTGAATGTAGTATTTGG	Myasthenic syndrome, congenital, without tubular aggregates
28940588	ALG3	NM_005787.5(ALG3):c.353G>A (p.Gly118Asp)	CCGAGRCACTGACATCCGCATGG	Congenital disorder of glycosylation type 1D
387906273	ALG3	NM_005787.5(ALG3):c.165C>T (p.Gly55=)	CCTGGCGGAGGTGGGYATCACCT	Congenital disorder of glycosylation type 1D
121908294	ALG8	NM_024079.4(ALG8):c.824G>A (p.Gly275Asp)	GGGRCCTCTGTCATGCATATTGG	Congenital disorder of glycosylation type 1H
397514527	ALOX12B	NM_001139.2(ALOX12B):c.1294C>T (p.Arg432Ter)	CCTCATCCCCCATACCYGATACA	Autosomal recessive congenital ichthyosis 2

397514531	ALOX12B	NM_001139.2(ALOX12B):c.1207C>T (p.His403Tyr)	CCCACCTGCTGGAGACAYACCTC, CCACCTGCTGGAGACAYACCTCA	Autosomal recessive congenital ichthyosis 2
121434233	ALOXE3	NM_001165960.1(ALOXE3):c.1096C>T (p.Arg366Ter)	CCTTGGGAATGAAGCTTYGAGGG	Autosomal recessive congenital ichthyosis 3
749544042	ALPL	NM_000478.4(ALPL):c.648+1G>A	ACATTGACRTGAGTGCTCGGGGG	
121918007	ALPL	NM_000478.4(ALPL):c.571G>A (p.Glu191Lys)	AGACAACRAGATGCCCCCTGAGG	Childhood hypophosphatasia, Infantile hypophosphatasia, Adult hypophosphatasia
121918009	ALPL	NM_000478.4(ALPL):c.1001G>A (p.Gly334Asp)	AAGGAGRCAGAATTGACCACGGG, CAAGGAGRCAGAATTGACCACGG	Infantile hypophosphatasia
121918013	ALPL	NM_000478.4(ALPL):c.346G>A (p.Ala116Thr)	CGCCACCRCCTACCTGTGTGGGG, CCGCCACCRCCTACCTGTGTGGG	Childhood hypophosphatasia, Infantile hypophosphatasia, Adult hypophosphatasia, Odontohypophosphatasia
121918019	ALPL	NM_000478.4(ALPL):c.526G>A (p.Ala176Thr)	CAGCRCCGCCTACGCCCACTCGG	Childhood hypophosphatasia, Infantile hypophosphatasia
121918015	ALPL	NM_000478.4(ALPL):c.323C>T (p.Pro108Leu)	CCAATGCCCAGGTCCYTGACAGC	Odontohypophosphatasia
121918020	ALPL	NM_000478.4(ALPL):c.814C>T (p.Arg272Cys)	CCCACTTCATCTGGAACYGCACG, CCACTTCATCTGGAACYGCACGG	Childhood hypophosphatasia, Infantile hypophosphatasia
1130335	ALPP	NM_001632.3(ALPP):c.74C= (p.Pro25=)	CCCTGGGCATCATCCYAGGTAAT, CCTGGGCATCATCCYAGGTAATG	
587776684	ALX1	NM_006982.2(ALX1):c.531+1G>A	TCCAGRTAGGAGCCAAAAAGAGG	Frontonasal dysplasia 3
121908168	ALX3	NM_006492.2(ALX3):c.547C>T (p.Arg183Trp)	CCTGATGTGTATGCCYGGGAGCA	Frontonasal dysplasia 1
387907269	AMER1	NM_152424.3(AMER1):c.811C>T (p.Gln271Ter)	CCTCAGCACATGTGYAACCCAAG	Osteopathia striata with cranial sclerosis
121912682	AMPD1	NM_000036.2(AMPD1):c.1373G>A (p.Arg458His)	GAGCCCCRCCTGTCCATCTATGG	Muscle AMP deaminase deficiency, not provided
121964981	AMT	NM_000481.3(AMT):c.806G>A (p.Gly269Asp)	AGGCAGRCCTCTGCCTGTATGGG, GAGGCAGRCCTCTGCCTGTATGG	Non-ketotic hyperglycinemia
36210415	ANK2	NM_001127493.1(ANK2):c.1360G>A (p.Gly454Arg)	CCGAGCCRGGCAGGTGGAAGTGG	Torsades de pointes
1800497	ANKK1	NM_178510.1(ANKK1):c.2137G>A (p.Glu713Lys)	CCAGCTGGGCGCCTGCCTYGACC	Dopamine receptor d2, reduced brain density of
137852512	ANOS1	NM_000216.2(ANOS1):c.711G>A (p.Trp237Ter)	TCACTGRCAGACAGTGGCCCAGG	Kallmann syndrome 1

137852514	ANOS1	NM_000216.2(ANOS1):c.774G>A (p.Trp258Ter)	CCGATGRTACCAGTTTCGAGTGG	Kallmann syndrome 1
137852516	ANOS1	NM_000216.2(ANOS1):c.784C>T (p.Arg262Ter)	CCGATGGTACCAGTTTYGAGTGG	Kallmann syndrome 1
137852517	ANOS1	NM_000216.2(ANOS1):c.1187C>T (p.Ser396Leu)	CCCTTCACTTCACATYGACACAT, CCTTCACTTCACATYGACACATG	Kallmann syndrome 1
397514700	ANTXR1	NM_032208.2(ANTXR1):c.505C>T (p.Arg169Ter)	CCCAGGCTAATAGGTCTYGAGAT, CCAGGCTAATAGGTCTYGAGATC	Odontotrichomelic syndrome
587776739	AP1S2	NM_003916.4(AP1S2):c.288+5G>A	CCTAAAATAAAATACTAYTCACA	Mental retardation X-linked syndromic 5
397514498	AP2S1	NM_004069.4(AP2S1):c.43C>T (p.Arg15Cys)	CCGGGCAGGCAAGACGYGCCTGG	Hypocalciuric hypercalcemia, familial, type 3
730882249	AP4M1	NM_004722.3(AP4M1):c.952C>T (p.Arg318Ter)	CCAGGTTTATCTAAAGTTGYGAT	Microcephaly, Hypoplasia of the corpus callosum, Spastic paraplegia 50, autosomal recessive, Global developmental delay, CNS hypomyelination, Brain atrophy
587781392	APC	NM_000038.5(APC):c.637C>T (p.Arg213Ter)	CCAGGATATGGAAAAAYGAGCAC	Hereditary cancer-predisposing syndrome, not provided
121913327	APC	NM_000038.5(APC):c.4012C>T (p.Gln1338Ter)	CCAAATCCAGCAGACTGYAGGGT	Familial adenomatous polyposis 1, Carcinoma of colon
587783029	APC	NM_000038.5(APC):c.3286C>T (p.Gln1096Ter)	CCAACCACATTTTGGACAGYAGG, CCACATTTTGGACAGYAGGAATG	Familial adenomatous polyposis 1, not provided
121909576	APOA4	NM_000482.3(APOA4):c.748G>A (p.Glu250Lys)	ACGCCRAGGAGCTCAAGGCCAGG	
121918390	APOB	NM_000384.2(APOB):c.7564C>T (p.Arg2522Ter)	CCTAGAAGATACACGAGACYGAA	Hypobetalipoproteinemia, familial, associated with apob32
138326449	APOC3	NM_000040.1(APOC3):c.55+1G>A	TCTGCCCRTAAGCACTTGGTGGG, CTCTGCCCRTAAGCACTTGGTGG	Coronary heart disease, Hyperalphalipoproteinemia 2
28931577	APOE	NM_000041.3(APOE):c.349G>A (p.Ala117Thr)	CAGRCGGCGCAGGCCCGGCTGGG, GCAGRCGGCGCAGGCCCGGCTGG, AGCTGCAGRCGGCGCAGGCCCGG	
121918398	APOE	NM_000041.3(APOE):c.875G>A (p.Arg292His)	GCAGCRCCAGTGGGCCGGGCTGG	
267606664	APOE	NM_000041.3(APOE):c.434G>A (p.Gly145Asp)	CATGCTCGRCCAGAGCACCGAGG	
7412	APOE	NM_000041.2(APOE):c.52	CCGATGACCTGCAGAAGYGCCTG	Familial type 3 hyperlipoproteinemia

		6C>T (p.Arg176Cys)		
769455	APOE	NM_000041.3(APOE):c.487C>T (p.Arg163Cys)	CCCACCTGCGCAAGCTGYGTAAG, CCACCTGCGCAAGCTGYGTAAGC	Familial type 3 hyperlipoproteinemia
387906567	APOE	NM_000041.3(APOE):c.478C>T (p.Arg160Cys)	CCTCGCCTCCCACCTGYGCAAGC	Familial type 3 hyperlipoproteinemia
63749810	APP	NM_000484.3(APP):c.2080G>A (p.Asp694Asn)	GAARATGTGGGTTCAAACAAAGG	Cerebral amyloid angiopathy, APP-related, not provided
63750734	APP	NM_000484.3(APP):c.2143G>A (p.Val715Met)	GACARTGATCGTCATCACCTTGG	Alzheimer disease, type 1, not provided
104894507	APRT	NM_000485.2(APRT):c.294G>A (p.Trp98Ter)	TCTGTGRGCCTCCTATTCCCTGG	Adenine phosphoribosyltransferase deficiency
121908131	APTX	NM_175073.2(APTX):c.617C>T (p.Pro206Leu)	CCATTGGCTGGTCTTACYGTGGA	Adult onset ataxia with oculomotor apraxia
104894004	AQP1	NM_198098.2(AQP1):c.113C>T (p.Pro38Leu)	CCCTGGGCTTCAAATACCYGGTG, CCTGGGCTTCAAATACCYGGTGG	
368292687	AQP5	NM_001651.3(AQP5):c.562C>T (p.Arg188Cys)	CCATGAACCCAGCCYGCTCTTTT	Diffuse palmoplantar keratoderma, Bothnian type
104894742	AR	NM_000044.3(AR):c.4G>A (p.Glu2Lys)	ATGRAAGTGCAGTTAGGGCTGGG, GATGRAAGTGCAGTTAGGGCTGG, CAAGGATGRAAGTGCAGTTAGGG	Reifenstein syndrome
137852563	AR	NM_000044.3(AR):c.2157G>A (p.Trp719Ter)	AAGTGRGCCAAGGCCTTGCCTGG	
137852571	AR	NM_000044.3(AR):c.2191G>A (p.Val731Met)	CTTACACRTGGACGACCAGATGG	Malignant tumor of prostate
137852572	AR	NM_000044.3(AR):c.2324G>A (p.Arg775His)	GGTACCRCATGCACAAGTCCCGG	
137852583	AR	NM_000044.3(AR):c.2164G>A (p.Ala722Thr)	GGCCAAGRCCTTGCCTGGTAAGG	Malignant tumor of prostate
137852588	AR	NM_000044.3(AR):c.1645C>T (p.Pro549Ser)	CCAGGGACCATGTTTTGYCCATT	Hypospadias 1, X-linked
28940281	ARHGEF10	NM_014629.3(ARHGEF10):c.995C>T (p.Thr332Ile)	CCGCGAAGGACGGCAYCAAGGAC	Slowed nerve conduction velocity, autosomal dominant
587779745	ARID1B	NM_020732.3(ARID1B):c.4102C>T (p.Gln1368Ter)	CCAGCCCGGCCTGTACCCAYAGC, CCCGGCCTGTACCCAYAGCAGCC,	Coffin Siris/Intellectual Disability

			CCGGCCTGTACCCAYAGCAGCCG	
387907140	ARID1B	NM_020732.3(ARID1B):c.3919C>T (p.Gln1307Ter)	CCCAACAGCAGCATGYAGGACAT, CCAACAGCAGCATGYAGGACATG	Mental retardation, autosomal dominant 12
387907141	ARID1B	NM_020732.3(ARID1B):c.3304C>T (p.Arg1102Ter)	CCCCTGGACCTGTTCYGACTCTA, CCCTGGACCTGTTCYGACTCTAC	Mental retardation, autosomal dominant 12
369721476	ARMC5	NM_001288767.1(ARMC5):c.1084C>T (p.Arg362Ter)	CCCTCCTGGAACTCAGCYGAGGC, CCTCCTGGAACTCAGCYGAGGCT	Acth-independent macronodular adrenal hyperplasia 2
199476366	ARSA	NM_000487.5(ARSA):c.737G>A (p.Arg246His)	TTCAGGCCRCGGGCCATTTGGGG	Metachromatic leukodystrophy, not provided
74315461	ARSA	NM_000487.5(ARSA):c.370G>A (p.Gly124Ser)	GGCCRGCAAGTGGCACCTTGGGG, TGGCCRGCAAGTGGCACCTTGGG, ATGGCCRGCAAGTGGCACCTTGG	Metachromatic leukodystrophy, not provided
80338815	ARSA	NM_000487.5(ARSA):c.465+1G>A	CGACCAGRTAGGAACCACCCGGG, ACGACCAGRTAGGAACCACCCGG	Metachromatic leukodystrophy, Metachromatic leukodystrophy, juvenile type, Metachromatic leukodystrophy, adult type
80338820	ARSA	NM_000487.5(ARSA):c.1210+1G>A	CCCAGGRTAACCCCTCCCCGTGG	Metachromatic leukodystrophy, Metachromatic leukodystrophy, juvenile type
74315458	ARSA	NM_000487.5(ARSA):c.257G>A (p.Arg86Gln)	GTTCRGATGGGCATGTACCCTGG	Metachromatic leukodystrophy
74315483	ARSA	NM_000487.5(ARSA):c.763G>A (p.Glu255Lys)	ATGRAGCTGGATGCAGCTGTGGG, GATGRAGCTGGATGCAGCTGTGG	Metachromatic leukodystrophy, Metachromatic leukodystrophy, late infantile
74315456	ARSA	NM_000487.5(ARSA):c.293C>T (p.Ser98Phe)	CCTGGTGCCCAGCTYCCGGGGGG	Metachromatic leukodystrophy, Metachromatic leukodystrophy, late infantile
74315462	ARSA	NM_000487.5(ARSA):c.413C>T (p.Pro138Leu)	CCTGAGGGGGCCTTCCTGCYCCC	
398123418	ARSA	NM_000487.5(ARSA):c.986C>T (p.Thr329Ile)	CCTGCCCCCAGGCGTGAYCCACG	Metachromatic leukodystrophy
74315468	ARSA	NM_000487.5(ARSA):c.677C>T (p.Ala226Val)	CCCTTCTTCCTGTACTATGYCTC, CCTTCTTCCTGTACTATGYCTCT	Metachromatic leukodystrophy
74315473	ARSA	NM_000487.5(ARSA):c.868C>T (p.Arg290Cys)	CCCAGACCTGAGACCATGYGTAT, CCAGACCTGAGACCATGYGTATG	Metachromatic leukodystrophy
74315481	ARSA	NM_000487.5(ARSA):c.12	CCCACAGTGATACCAYTGCAGAC	Metachromatic leukodystrophy

		32C>T (p.Thr411Ile)	, CCACAGTGATACCAYTGCAGACC	
118203941	ARSB	NM_000046.3(ARSB):c.1214G>A (p.Cys405Tyr)	CCCAGRTCCCAGGAACAGCATGG	Mucopolysaccharidosis type VI, MUCOPOLYSACCHARIDOSIS, TYPE VI, SEVERE
118203942	ARSB	NM_000046.3(ARSB):c.284G>A (p.Arg95Gln)	TCGCRGAGCCAGCTGCTCACTGG	Mucopolysaccharidosis type VI, not provided
122460155	ARSE	NM_000047.2(ARSE):c.1475G>A (p.Cys492Tyr)	CGGTGCCTRCTATGGAAGAAAGG	Chondrodysplasia punctata 1, X-linked recessive
28935474	ARSE	NM_000047.2(ARSE):c.1732C>T (p.Pro578Ser)	CCCTGCTGTGGCCCGTTCYCCCT, CCTGCTGTGGCCCGTTCYCCCTC	Chondrodysplasia punctata 1, X-linked recessive
587783096	ARX	NM_139058.2(ARX):c.1141G>A (p.Ala381Thr)	TCGGRCCAAGTGGCGCAAGCGGG , GTCGGRCCAAGTGGCGCAAGCGG	not provided
104894740	ARX	NM_139058.2(ARX):c.1117C>T (p.Gln373Ter)	CCGAGGCCCGAGTCYAGGTGAGC	Lissencephaly 2, X-linked
104894743	ARX	NM_139058.2(ARX):c.1058C>T (p.Pro353Leu)	CCAGAAGACGCACTACCYGGACG	Epileptic encephalopathy, early infantile, 1
587783189	ARX	NM_139058.2(ARX):c.1414C>T (p.Arg472Ter)	CCTCGGAGCGGCAGTGTTCYGAC	Lissencephaly 2, X-linked
145873635	ASAH1	NM_004315.4(ASAH1):c.173C>T (p.Thr58Met)	ACTCACRTTGGTCCTGAAGGAGG	Jankovic Rivera syndrome
104886478	ASB10	NM_080871.3(ASB10):c.765C>T (p.Thr255=)	CCGATGCCGAGGCCACCACYGCC	Glaucoma 1, open angle, F
145138923	ASL	NM_000048.3(ASL):c.35G>A (p.Arg12Gln)	TGGCCRGTTTGTGGGTGCAGTGG	Argininosuccinate lyase deficiency, not provided
142637046	ASL	NM_000048.3(ASL):c.446+1G>A	GCAGAGGCRTGAGTCCTACAGGG	not provided
28940286	ASL	NM_001024943.1(ASL):c.1153C>T (p.Arg385Cys)	CCCAGATGCCATTCYGCCAGGCC	Argininosuccinate lyase deficiency
398123126	ASL	NM_000048.3(ASL):c.544C>T (p.Arg182Ter)	CCACGCCGTGGCACTGACCYGAG , CCGTGGCACTGACCYGAGACTCT	Argininosuccinate lyase deficiency, not provided
374304304	ASL	NM_000048.3(ASL):c.280C>T (p.Arg94Cys)	CCACACAGCCAATGAGYGCCGCC	Argininosuccinate lyase deficiency, not provided
398122974	ASNS	NM_183356.3(ASNS):c.1648C>T (p.Arg550Cys)	CCACTGACCCTTCTGCCYGCACG	Asparagine synthetase deficiency, Abnormality of neuronal migration
199422154	ASPM	NM_018136.4(ASPM):c.3082G>A (p.Gly1028Arg)	GCATRGTAAAAACTGAGTAGAGG	Primary autosomal recessive microcephaly 5
587783287	ASPM	NM_018136.4(ASPM):c.90	CCTTATAGAGACATYGAGCTGCT	Primary autosomal recessive microcephaly 5

		91C>T (p.Arg3031Ter)		
587783227	ASPM	NM_018136.4(ASPM):c.2791C>T (p.Arg931Ter)	CCTTTTGGCTTTTTCAYGAGATT	Primary autosomal recessive microcephaly 5
587783275	ASPM	NM_018136.4(ASPM):c.8017C>T (p.Gln2673Ter)	CCAAGCAGTTATTTGTATAYAGT	Primary autosomal recessive microcephaly 5
199422148	ASPM	NM_018136.4(ASPM):c.1990C>T (p.Gln664Ter)	CCCATTATCGCTGTGGCAYAGTC, CCATTATCGCTGTGGCAYAGTCC	Primary autosomal recessive microcephaly 5
199422175	ASPM	NM_018136.4(ASPM):c.7894C>T (p.Gln2632Ter)	CCAGGCTGCCATTATTATTYAGA	Primary autosomal recessive microcephaly 5
137852996	ASPM	NM_018136.4(ASPM):c.349C>T (p.Arg117Ter)	CCACTCAAAGAAGGCYGAGTAAG	Primary autosomal recessive microcephaly 5
121908637	ASS1	NM_000050.4(ASS1):c.470G>A (p.Arg157His)	CAAGGGCCRCAATGACCTGATGG	Citrullinemia type I, not provided
121908639	ASS1	NM_000050.4(ASS1):c.970G>A (p.Gly324Ser)	TACCRGTGCGTAAGACTCTATGG	Citrullinemia type I, not provided
777828000	ASS1	NM_000050.4(ASS1):c.571G>A (p.Glu191Lys)	CAGCTACRAGGCTGGAATCCTGG	Citrullinemia type I
398123131	ASS1	NM_000050.4(ASS1):c.794G>A (p.Arg265His)	CGTGGGCCRTATTGACATCGTGG	Citrullinemia type I, not provided
786204537	ASS1	NM_000050.4(ASS1):c.1030C>T (p.Arg344Ter)	CCAAGTCCCAGGAGYGAGTGGAA	Citrullinemia type I, not provided
138350285	ASS1	NM_000050.4(ASS1):c.-4C>T	CCTCGACTCCCGCCAGAYGCTAT	not provided
121908640	ASS1	NM_000050.4(ASS1):c.1087C>T (p.Arg363Trp)	CCAGGTGTACATCCTCGGCYGGG	Citrullinemia type I, not provided
121908642	ASS1	NM_000050.4(ASS1):c.910C>T (p.Arg304Trp)	CCTTCACCATGGACYGGGAAGTG	Citrullinemia type I
373145711	ASXL1	NM_015338.5(ASXL1):c.1210C>T (p.Arg404Ter)	CCAGCCACCCGACAGYGAGATGG	C-like syndrome
387907077	ASXL1	NM_015338.5(ASXL1):c.2773C>T (p.Gln925Ter)	CCATCTGTTGAGCCCYAGGTTGG	C-like syndrome
587777061	ASXL3	NM_030632.1(ASXL3):c.1210C>T (p.Gln404Ter)	CCTTGGCAGAACAAYAGCCAAAA	Bainbridge-Ropers syndrome
761357250	ATF6	NM_007348.3(ATF6):c.970C>T (p.Arg324Cys)	CCGCTTGTCAGTCTYGCAAGAAG	Achromatopsia 7
119476046	ATL1	NM_015915.4(ATL1):c.715C>T (p.Arg239Cys)	CCAAATTCTTGGAAAAAYGCCTC	Spastic paraplegia 3
137852657	ATL1	NM_015915.4(ATL1):c.467C>T (p.Thr156Ile)	CCTTTGATAGTCAGTCAAYTTTG	Spastic paraplegia 3

587779818	ATM	NM_000051.3(ATM):c.170G>A (p.Trp57Ter)	TGAATTRGGATGCTGTTTTTAGG	Ataxia-telangiectasia syndrome, Hereditary cancer-predisposing syndrome
786201957	ATM	NM_000051.3(ATM):c.3349C>T (p.Gln1117Ter)	CCTTTGAAGCTTCAGYAAACAGC	Hereditary cancer-predisposing syndrome
199624796	ATP13A2	NM_022089.3(ATP13A2):c.490C>T (p.Arg164Trp)	ATACCRCAGCACCCGCTTCTGGG, AATACCRCAGCACCCGCTTCTGG	Parkinson disease 9
121918616	ATP1A2	NM_000702.3(ATP1A2):c.1643G>A (p.Arg548His)	GGGAGCRTGTGCTGGGTGAGAGG	Migraine, familial basilar
796052276	ATP1A2	NM_000702.3(ATP1A2):c.1091C>T (p.Thr364Met)	CCTGGAGGCGGTGGAGAYGCTGG	not provided
121918615	ATP1A2	NM_000702.3(ATP1A2):c.2936C>T (p.Pro979Leu)	CCCTCCGCATGTACCYGCTCAAG, CCTCCGCATGTACCYGCTCAAGT	Familial hemiplegic migraine type 2
121918620	ATP1A2	NM_000702.3(ATP1A2):c.1127C>T (p.Thr376Met)	CCATCTGCTCGGACAAGAYGGGC	Familial hemiplegic migraine type 2
80356534	ATP1A3	NM_152296.4(ATP1A3):c.1838C>T (p.Thr613Met)	CCGGCGATCACCCCATCAYGGCC	Dystonia 12
121918113	ATP2A1	NM_004320.4(ATP2A1):c.592C>T (p.Arg198Ter)	CCCGTTCCTGACCCCYGAGCTGT, CCGTTCCTGACCCCYGAGCTGTC	Brody myopathy
121918115	ATP2A1	NM_004320.4(ATP2A1):c.2366C>T (p.Pro789Leu)	CCTGAGGCCCTGATCCYGGTGCA	Brody myopathy
121912736	ATP2A2	NM_001681.3(ATP2A2):c.2305G>A (p.Gly769Arg)	CGTCRGGGAAGTTGTCTGGTAGG, CCAACGTCRGGGAAGTTGTCTGG	Darier disease, segmental
28929478	ATP2A2	NM_001681.3(ATP2A2):c.68G>A (p.Gly23Glu)	TACGGRGCTGAGCCTGGAACAGG	Keratosis follicularis
137853012	ATP2C1	NM_001001486.1(ATP2C1):c.910G>A (p.Ala304Thr)	GCTGTARCAGCAATTCCTGAAGG	Familial benign pemphigus
121918521	ATP6AP2	NM_005765.2(ATP6AP2):c.321C>T (p.Asp107=)	CCTTTTAGTCTTGAYAGTGTTGC	Mental retardation, X-linked, syndromic, Hedera type
374480381	ATP6V0A2	NM_012463.3(ATP6V0A2):c.1514+1G>A	GGAARTAAGTGTCCCATAGCTGG	Cutis laxa with osteodystrophy, not provided
28939081	ATP6V0A4	NM_020632.2(ATP6V0A4):c.2420G>A (p.Arg807Gln)	TGCRACTGCACTGGTAAGGATGG, GCCCTGCRACTGCACTGGTAAGG	Renal tubular acidosis, distal, autosomal recessive, with late-onset sensorineural hearing loss
121908367	ATP6V0A4	NM_020632.2(ATP6V0A4):c.2257C>T (p.Gln753Ter)	CCTCAGCCTGGCTCATGCAYGTG	Renal tubular acidosis, distal, autosomal recessive
121908368	ATP6V0A4	NM_020632.2(ATP6V0A4):c.1571C>T (p.Pro524Leu)	CCCGTTTGGGATTGATCYGGTAA, CCGTTTGGGATTGATCYGGTAAT	Renal tubular acidosis, distal, autosomal recessive
121964881	ATP6V1B1	NM_001692.3(ATP6V1B1):c.232G>A (p.Gly78Arg)	GAGCRGGCAGGTGCTTGAGGTGG,	

			GAGGAGCRGGCAGGTGCTTGAG G	
794729667	ATP6V1B 2	NM_001693.3(ATP6V1B2) :c.1516C>T (p.Arg506Ter)	CCTCAGCGAATTTTACCCTYGAG	Zimmermann-Laband syndrome 2
267606673	ATP7A	NM_000052.6(ATP7A):c.2 981C>T (p.Thr994Ile)	CCAAGCCTCTATCAYAGTTCTGT	Distal spinal muscular atrophy, X-linked 3
201038679	ATP7B	NM_000053.3(ATP7B):c.2 975C>T (p.Pro992Leu)	CGTGRGCGTGGCCAGCCCCAGGG , CCGTGRGCGTGGCCAGCCCCAGG	Wilson disease
28942076	ATP7B	NM_000053.3(ATP7B):c.2 827G>A (p.Gly943Ser)	GTAATCRGTTTTATCGATTTTGG	Wilson disease
137853283	ATP7B	NM_000053.3(ATP7B):c.2 336G>A (p.Trp779Ter)	GGGCCGGTRGCTGGAACACTTGG	Wilson disease
587783306	ATP7B	NM_000053.3(ATP7B):c.2 865+1G>A	TCCTRTAAGTTGAATGCCTTGGG, TTCCTRTAAGTTGAATGCCTTGG	Wilson disease
72552255	ATP7B	NM_000053.3(ATP7B):c.2 930C>T (p.Thr977Met)	CCGGAACCCAAGTTCRTCACGTT	Wilson disease, not provided
121907994	ATP7B	NM_000053.3(ATP7B):c.2 621C>T (p.Ala874Val)	CCCGGAAGCACTGTAATTGYGGG , CCGGAAGCACTGTAATTGYGGGG	Wilson disease
121909101	ATP8B1	NM_005603.4(ATP8B1):c. 1660G>A (p.Asp554Asn)	CTCTCCCRATGAAGGTGCCCTGG	Progressive intrahepatic cholestasis
122445104	ATRX	NM_000489.4(ATRX):c.52 25G>A (p.Arg1742Lys)	AGGARGAGGATTATTTTAACAGG	ATR-X syndrome
122445099	ATRX	NM_000489.4(ATRX):c.71 56C>T (p.Arg2386Ter)	CCAGGAGCTTGATGTTAAAYGAA	ATR-X syndrome
730880309	AUH	NM_001698.2(AUH):c.895- 1G>A	CTCARGTCGATTTAGTAACAGGG, TCTCARGTCGATTTAGTAACAGG	3-Methylglutaconic aciduria
121908654	AURKC	NM_001015879.1(AURKC):c.629G>A (p.Cys210Tyr)	CTCTRCTATGAGCTGCTGGTGGG, GCTCTRCTATGAGCTGCTGGTGG, AGTGCTCTRCTATGAGCTGCTGG	Infertility associated with multi-tailed spermatozoa and excessive DNA
121964882	AVP	NM_000490.4(AVP):c.262 G>A (p.Gly88Ser)	AGTCCRGCCAGAAGGCGTGCGGG , CAGTCCRGCCAGAAGGCGTGCGG	Neurohypophyseal diabetes insipidus
121964890	AVP	NM_000490.4(AVP):c.260 C>T (p.Ser87Phe)	CCGTCGCCCTGCCAGTYCGGCCA	Neurohypophyseal diabetes insipidus
121964892	AVP	NM_000490.4(AVP):c.20C >T (p.Pro7Leu)	CCTGACACCATGCTGCYCGCCTG	
28935496	AVPR2	NM_000054.4(AVPR2):c.3	CCAGATGCCCTGTGTYGGGCCGT	Nephrogenic diabetes insipidus, X-linked

		37C>T (p.Arg113Trp)		
104894760	AVPR2	NM_000054.4(AVPR2):c.310C>T (p.Arg104Cys)	CCTGGAAGGCCACCGACYGCTTC	Nephrogenic diabetes insipidus, Nephrogenic diabetes insipidus, X-linked
730882193	AXIN2	NM_004655.3(AXIN2):c.1989G>A (p.Trp663Ter)	ATCTGTGRGGGGGCAACAGCGGG, CATCTGTGRGGGGGCAACAGCGG	Oligodontia-colorectal cancer syndrome
121908568	AXIN2	NM_004655.3(AXIN2):c.1966C>T (p.Arg656Ter)	CCGCTCGTCTCCAGGCGAAYGAG	Oligodontia-colorectal cancer syndrome
367543074	B3GALNT2	NM_152490.4(B3GALNT2):c.802G>A (p.Val268Met)	GGTRTGGAGGGAGTTGCAGGTGG, GAAGGTRTGGAGGGAGTTGCAGG	Congenital muscular dystrophy-dystroglycanopathy with brain and eye anomalies, type A11, not provided
397514722	B3GALT6	NM_080605.3(B3GALT6):c.16C>T (p.Arg6Trp)	CGCCCRCCGCAGCAGCTTCATGG	Ehlers-Danlos syndrome, progeroid type, 2
397514724	B3GALT6	NM_080605.3(B3GALT6):c.649G>A (p.Gly217Ser)	CCGAGAGCACGTAGCYGCCGCCC	
387906876	BAG3	NM_004281.3(BAG3):c.1430G>A (p.Arg477His)	GTGCRTCAGGCCAGGAGAGACGG	Dilated cardiomyopathy 1HH
387906874	BAG3	NM_004281.3(BAG3):c.211C>T (p.Arg71Trp)	CCAATGGCCCTTCCYGGGAGGGC	Dilated cardiomyopathy 1HH, not provided
387906871	BANF1	NM_003860.3(BANF1):c.34G>A (p.Ala12Thr)	CTTCGTGRCAGAGCCCATGGGGG, ACTTCGTGRCAGAGCCCATGGGG	Nestor-Guillermo progeria syndrome
786202118	BARD1	NM_000465.3(BARD1):c.2268G>A (p.Trp756Ter)	TCTGRAAGGCTCCTTCGAGCTGG	Hereditary cancer-predisposing syndrome
587777829	BBS1	NM_024649.4(BBS1):c.432+1G>A	AAGAGRTAAATAAATAACATGGG, AAAGAGRTAAATAAATAACATGG	Bardet-Biedl syndrome, Bardet-Biedl syndrome 1
121908178	BBS2	NM_031885.3(BBS2):c.943C>T (p.Arg315Trp)	CCTTCTGTTCAGTCYGGGGCTAC	
121908180	BBS2	NM_031885.3(BBS2):c.646C>T (p.Arg216Ter)	CCCATGTATGGCAGTYGATTTGG, CCATGTATGGCAGTYGATTTGGT	
119466002	BBS7	NM_176824.2(BBS7):c.632C>T (p.Thr211Ile)	CCTTTTGTTTGGGAYATCAGACG	Bardet-Biedl syndrome
121918133	BCAM	NM_005581.4(BCAM):c.361C>T (p.Arg121Ter)	CCCAGGTGGGCGACGAGYGAGAC, CCAGGTGGGCGACGAGYGAGACT	

375785084	BCKDHA	NM_000709.3(BCKDHA):c.659C>T (p.Ala220Val)	CCTGCAGCGGTGGGGGCGGYGTA	Maple syrup urine disease, not provided
398123497	BCKDHA	NM_000709.3(BCKDHA):c.288+9C>T	CCCCCACGTGAGAGGYGGCCTCC, CCCCACGTGAGAGGYGGCCTCCC	Maple syrup urine disease, not provided
398123503	BCKDHA	NM_000709.3(BCKDHA):c.632C>T (p.Thr211Met)	CCTCTCCACTGGCCAYGCAGATC	Maple syrup urine disease, not provided
137852873	BCKDHA	NM_000709.3(BCKDHA):c.793C>T (p.Arg265Trp)	CCCCATCATCTTCTTCTGCYGGA, CCCATCATCTTCTTCTGCYGGAA, CCATCATCTTCTTCTGCYGGAAC	Maple syrup urine disease type 1A
398124602	BCKDHB	NM_000056.3(BCKDHB):c.952-1G>A	TTCARTCTGTGATCAAAACAGGG, TTTCARTCTGTGATCAAAACAGG	Maple syrup urine disease, not provided
121965004	BCKDHB	NM_000056.3(BCKDHB):c.616C>T (p.His206Tyr)	CCTGAAGCATTTTTTGCCYATTG	
397514573	BCKDK	NM_005881.3(BCKDK):c.466C>T (p.Arg156Ter)	CCAGTACTGCCAGCTGGTGYGAC	Branched-chain ketoacid dehydrogenase kinase deficiency
121908571	BCS1L	NM_004328.4(BCS1L):c.830G>A (p.Ser277Asn)	GCAGARCCTGGTACTCCTGGAGG, GCAGCAGARCCTGGTACTCCTGG	Mitochondrial complex III deficiency
121908578	BCS1L	NM_004328.4(BCS1L):c.550C>T (p.Arg184Cys)	CCCTTTGGCTATCCACGCYGCCG, CCTTTGGCTATCCACGCYGCCGG	Mitochondrial complex III deficiency
28940276	BEST1	NM_004183.3(BEST1):c.25G>A (p.Val9Met)	CAARTGGCTAATGCCCGCTTAGG	Vitelliform dystrophy, not provided
121918287	BEST1	NM_004183.3(BEST1):c.949G>A (p.Val317Met)	CCAGRTGTCCCTGTTGGCTGTGG	Bestrophinopathy, autosomal recessive
28940570	BEST1	NM_004183.3(BEST1):c.728C>T (p.Ala243Val)	CCCAGGTGGTGACTGTGGYGGTG, CCAGGTGGTGACTGTGGYGGTGT	Vitelliform dystrophy, not provided
372989281	BEST1	NM_004183.3(BEST1):c.763C>T (p.Arg255Trp)	CCTGACTTGTCTAGTTGGGYGGC	Retinitis pigmentosa
398123028	BICD2	NM_015250.3(BICD2):c.320C>T (p.Ser107Leu)	CCTGATCCAGGAGTYGGCCTCCA	Spinal muscular atrophy, lower extremity predominant 2, autosomal dominant
587783343	BIN1	NM_139343.2(BIN1):c.1713G>A (p.Trp571Ter)	CTGRAACCAGCACAAGGAGCTGG	Autosomal recessive centronuclear myopathy
55758736	BLK	NM_001715.2(BLK):c.211G>A (p.Ala71Thr)	CTACACCRCTATGAATGATCGGG, ACTACACCRCTATGAATGATCGG	Maturity-onset diabetes of the young, type 11, not specified
367543025	BLM	NM_000057.3(BLM):c.3197G>A (p.Cys1066Tyr)	TGATAATTRCTGTAAAACAAAGG	Bloom syndrome
104894763	BMP15	NM_005448.2(BMP15):c.2	CCTAGGGCATTCACTGYGGTACA	Premature ovarian failure 4

		02C>T (p.Arg68Trp)		
137853320	BMP15	NM_005448.2(BMP15):c.631C>T (p.Gln211Ter)	CCGTTTTATGTGTCAGYAGCAAA	Premature ovarian failure 4
121912766	BMP4	NM_001202.3(BMP4):c.1037C>T (p.Ala346Val)	CCCCTTTCCACTGGYTGACCACC	Orofacial cleft 11
199476088	BMPR1A	NM_004329.2(BMPR1A):c.1127G>A (p.Cys376Tyr)	AGTTRCTGCATTGCTGACCTGGG, GAGTTRCTGCATTGCTGACCTGG	Juvenile polyposis syndrome
764466442	BMPR1A	NM_004329.2(BMPR1A):c.1081C>T (p.Arg361Ter)	CCCGCAATTGCTCATYGAGACCT, CCGCAATTGCTCATYGAGACCTA	Hereditary cancer-predisposing syndrome
137852744	BMPR2	NM_001204.6(BMPR2):c.1040G>A (p.Cys347Tyr)	ACCTRTGTTATTAGTGACTTTGG	Primary pulmonary hypertension
137852746	BMPR2	NM_001204.6(BMPR2):c.1471C>T (p.Arg491Trp)	CCAGGATGCAGAGGCTYGGCTTA	Primary pulmonary hypertension
137852751	BMPR2	NM_001204.6(BMPR2):c.994C>T (p.Arg332Ter)	CCTGCAATTTCCCATYGAGATTT	Primary pulmonary hypertension
137852756	BMPR2	NM_001204.6(BMPR2):c.1297C>T (p.Gln433Ter)	CCGTACCAGAGTACYAGATGGCT	Pulmonary hypertension, primary, 1, with hereditary hemorrhagic telangiectasia
121964925	BPGM	NM_199186.2(BPGM):c.268C>T (p.Arg90Cys)	CCTGGCGTCTAAATGAGYGTCAC	Deficiency of bisphosphoglycerate mutase
397516894	BRAF	NM_004333.4(BRAF):c.1720C>T (p.His574Tyr)	CCAAGTCAATCATCYACAGAGAC	Cardiofaciocutaneous syndrome
397509284	BRCA1	NM_007294.3(BRCA1):c.5445G>A (p.Trp1815Ter)	GATGCCTGRACAGAGGACAATGG	Familial cancer of breast, Breast-ovarian cancer, familial 1, Hereditary cancer-predisposing syndrome
80356937	BRCA1	NM_007294.3(BRCA1):c.5212G>A (p.Gly1738Arg)	GTCAGARGAGATGTGGTCAATGG	Familial cancer of breast, Breast-ovarian cancer, familial 1
80356962	BRCA1	NM_007294.3(BRCA1):c.5444G>A (p.Trp1815Ter)	GATGCCTRGACAGAGGACAATGG	Familial cancer of breast, Hereditary breast and ovarian cancer syndrome, Breast-ovarian cancer, familial 1, Hereditary cancer-predisposing syndrome
80357219	BRCA1	NM_007294.3(BRCA1):c.5345G>A (p.Trp1782Ter)	GAATRGATGGTACAGCTGTGTGG	Familial cancer of breast, Breast-ovarian cancer, familial 1
80357284	BRCA1	NM_007294.3(BRCA1):c.5346G>A (p.Trp1782Ter)	GAATGRATGGTACAGCTGTGTGG	Familial cancer of breast, Breast-ovarian cancer, familial 1, Hereditary cancer-predisposing syndrome
80357292	BRCA1	NM_007294.3(BRCA1):c.962G>A (p.Trp321Ter)	TAACAGATRGGCTGGAAGTAAGG	Familial cancer of breast, Hereditary breast and ovarian cancer syndrome, Breast-ovarian cancer, familial 1, Hereditary cancer-predisposing syndrome
80358008	BRCA1	NM_007294.3(BRCA1):c.4676-1G>A	TTCARAGGGAACCCCTTACCTGG	Breast-ovarian cancer, familial 1
80358070	BRCA1	NM_007294.3(BRCA1):c.4097-1G>A	TTTAARGTGAAGCAGCATCTGGG, ATTTAARGTGAAGCAGCATCTGG	Familial cancer of breast, Breast-ovarian cancer, familial 1

62625307	BRCA1	NM_007294.3(BRCA1):c.3598C>T (p.Gln1200Ter)	CCCATACACATTTGGCTYAGGGT, CCATACACATTTGGCTYAGGGTT	Familial cancer of breast, Hereditary breast and ovarian cancer syndrome, Breast-ovarian cancer, familial 1, Hereditary cancer-predisposing syndrome
80356952	BRCA1	NM_007294.3(BRCA1):c.1630C>T (p.Gln544Ter)	CCAAACGGAGCAGAATGGTYAAG	Familial cancer of breast, Breast-ovarian cancer, familial 1
80357089	BRCA1	NM_007294.3(BRCA1):c.3331C>T (p.Gln1111Ter)	CCTGAAATAAAAAAGYAAGAATA	Familial cancer of breast, Breast-ovarian cancer, familial 1
80357123	BRCA1	NM_007294.3(BRCA1):c.5251C>T (p.Arg1751Ter)	CCACCAAGGTCCAAAGYGAGCAA	Familial cancer of breast, Hereditary breast and ovarian cancer syndrome, Breast-ovarian cancer, familial 1, Hereditary cancer-predisposing syndrome
80357211	BRCA1	NM_007294.3(BRCA1):c.949C>T (p.Gln317Ter)	CCTGGCTTAGCAAGGAGCYAACA	Familial cancer of breast, Breast-ovarian cancer, familial 1
80357372	BRCA1	NM_007294.3(BRCA1):c.415C>T (p.Gln139Ter)	CCGTGCCAAAAGACTTCTAYAGA, CCAAAAGACTTCTAYAGAGTGAA	Familial cancer of breast, Breast-ovarian cancer, familial 1
80357471	BRCA1	NM_007294.3(BRCA1):c.178C>T (p.Gln60Ter)	CCAGAAGAAAGGGCCTTCAYAGT	Familial cancer of breast, Breast-ovarian cancer, familial 1
587781506	BRCA2	NM_000059.3(BRCA2):c.7877G>A (p.Trp2626Ter)	TAGATRGATCATATGGAAACTGG	Hereditary breast and ovarian cancer syndrome, Hereditary cancer-predisposing syndrome
80358543	BRCA2	NM_000059.3(BRCA2):c.2978G>A (p.Trp993Ter)	AACAAATRGGCAGGACTCTTAGG	Breast-ovarian cancer, familial 2
80358544	BRCA2	NM_000059.3(BRCA2):c.2979G>A (p.Trp993Ter)	AACAAATGRGCAGGACTCTTAGG	Familial cancer of breast, Breast-ovarian cancer, familial 2, Hereditary cancer-predisposing syndrome
80359015	BRCA2	NM_000059.3(BRCA2):c.7886G>A (p.Trp2629Ter)	CATATRGAAACTGGCAGCTATGG	Familial cancer of breast, Breast-ovarian cancer, familial 2
80359205	BRCA2	NM_000059.3(BRCA2):c.9317G>A (p.Trp3106Ter)	GTTTTRGATAGACCTTAATGAGG	Familial cancer of breast, Breast-ovarian cancer, familial 2
80359803	BRCA2	NM_000059.3(BRCA2):c.8754G>A (p.Glu2918=)	CCTTGARGTGAGAGAGTAAGAGG	Familial cancer of breast, Breast-ovarian cancer, familial 2, Hereditary cancer-predisposing syndrome
730881581	BRCA2	NM_000059.3(BRCA2):c.8174G>A (p.Trp2725Ter)	AGATGGGTRGTATGCTGTTAAGG	Familial cancer of breast
276174913	BRCA2	NM_000059.3(BRCA2):c.8869C>T (p.Gln2957Ter)	CCATGGAATCTGCTGAAYAAAAG	Familial cancer of breast, Breast-ovarian cancer, familial 2
80358515	BRCA2	NM_000059.3(BRCA2):c.250C>T (p.Gln84Ter)	CCAATAATATTCAAAGAGYAAGG	Familial cancer of breast, Hereditary breast and ovarian cancer syndrome, Breast-ovarian cancer, familial 2
80358578	BRCA2	NM_000059.3(BRCA2):c.3319C>T (p.Gln1107Ter)	CCATAATTTAACACCTAGCYAAA	Familial cancer of breast, Breast-ovarian cancer, familial 2, Hereditary cancer-predisposing syndrome
80358851	BRCA2	NM_000059.3(BRCA2):c.6124C>T (p.Gln2042Ter)	CCAGAACATTTAATATCCYAAAA	Hereditary breast and ovarian cancer syndrome, Breast-ovarian cancer, familial 2

80358920	BRCA2	NM_000059.3(BRCA2):c.6952C>T (p.Arg2318Ter)	CCTAGGCACAATAAAAGATYGAA	Familial cancer of breast, Breast-ovarian cancer, familial 2
587782010	BRCA2	NM_000059.3(BRCA2):c.8608C>T (p.Gln2870Ter)	CCTTATTCACTAAAATTYAGGAG	Hereditary cancer-predisposing syndrome
587782613	BRCA2	NM_000059.3(BRCA2):c.3412C>T (p.Gln1138Ter)	CCAAGCTACATATTGYAGAAGAG	Hereditary cancer-predisposing syndrome
397507395	BRCA2	NM_000059.3(BRCA2):c.7963C>T (p.Gln2655Ter)	CCCAGAAAGGGTGCTTCTTYAAC, CCAGAAAGGGTGCTTCTTYAACT	Familial cancer of breast, Breast-ovarian cancer, familial 2
397507617	BRCA2	NM_000059.3(BRCA2):c.196C>T (p.Gln66Ter)	CCTATTTAAAACTCCAYAAAGGA	Familial cancer of breast, Breast-ovarian cancer, familial 2, Hereditary cancer-predisposing syndrome
587782539	BRIP1	NM_032043.2(BRIP1):c.2576-1G>A	TTCTTTARGACTTTCTAAATGGG, TTTCTTTARGACTTTCTAAATGG	Hereditary cancer-predisposing syndrome
137852985	BRIP1	NM_032043.2(BRIP1):c.897G>A (p.Met299Ile)	AGTGCATRGAATTGCTAGATGGG, AAGTGCATRGAATTGCTAGATGG	Breast cancer, early-onset
587782574	BRIP1	NM_032043.2(BRIP1):c.2377C>T (p.Gln793Ter)	CCAAATGTGAAAGATCTAYAGGT	Hereditary cancer-predisposing syndrome
730881633	BRIP1	NM_032043.2(BRIP1):c.1066C>T (p.Arg356Ter)	CCATATTACACAGCCYGAGAACT	Hereditary cancer-predisposing syndrome
74315287	BSND	NM_057176.2(BSND):c.28G>A (p.Gly10Ser)	GGATCRGCTTCATTGTGCTGGGG, CGGATCRGCTTCATTGTGCTGGG, CCGGATCRGCTTCATTGTGCTGG	Bartter syndrome type 4
74315289	BSND	NM_057176.2(BSND):c.139G>A (p.Gly47Arg)	TGGTGATCRGGGGCATCATCTGG	Bartter syndrome type 4
146015592	BTD	NM_000060.3(BTD):c.470G>A (p.Arg157His)	GCTCCAGCRCCTGAGTTGTATGG	Biotinidase deficiency, not provided
397514396	BTD	NM_000060.3(BTD):c.934G>A (p.Gly312Ser)	AAGTRGCATACACACCCCTCTGG	Biotinidase deficiency
397514343	BTD	NM_000060.3(BTD):c.236G>A (p.Arg79His)	CATCAGCCRCCAAGAGGCCTTGG	Biotinidase deficiency
397514375	BTD	NM_000060.3(BTD):c.595G>A (p.Val199Met)	AATGTCRTGTTCAGCAATAATGG	Biotinidase deficiency
397514417	BTD	NM_000060.3(BTD):c.1333G>A (p.Gly445Arg)	GATRGGCTTCACACAGTACATGG	Biotinidase deficiency
397514428	BTD	NM_000060.3(BTD):c.1610G>A (p.Gly537Glu)	CTATGRGCGCTTGTATGAGAGGG, TCTATGRGCGCTTGTATGAGAGG	Biotinidase deficiency
397514429	BTD	NM_000060.3(BTD):c.1613G>A (p.Arg538His)	CTATGGGCRCTTGTATGAGAGGG	Biotinidase deficiency
367902696	BTD	NM_000060.3(BTD):c.443	TCACCRCTTCAATGACACAGAGG	Biotinidase deficiency

		G>A (p.Arg148His)		
34885143	BTD	NM_000060.3(BTD):c.133 G>A (p.Gly45Arg)	CCACACCRGGGAGGAGAGCGTG G	Biotinidase deficiency, not specified, not provided
146600671	BTD	NM_000060.3(BTD):c.1369 G>A (p.Val457Met)	TCCAARTGTGTGCCCTGGTCAGG	Biotinidase deficiency
377651057	BTD	NM_000060.3(BTD):c.935 G>A (p.Gly312Asp)	AAGTGRCATACACACCCCTCTGG	Biotinidase deficiency
104893687	BTD	NM_000060.3(BTD):c.235 C>T (p.Arg79Cys)	CCCTCTGGCTCTCATCAGCYGCC, CCTCTGGCTCTCATCAGCYGCCA	Biotinidase deficiency, not provided
104893688	BTD	NM_000060.3(BTD):c.1595 C>T (p.Thr532Met)	CCTCTGGGCTGGTGAYGGCGGCT	Biotinidase deficiency, not provided
397514349	BTD	NM_000060.3(BTD):c.283 C>T (p.Gln95Ter)	CCTTGACATCTATGAAYAGCAAG	Biotinidase deficiency
397514363	BTD	NM_000060.3(BTD):c.469 C>T (p.Arg157Cys)	CCTCTAGGTGCTCCAGYGCCTGA	Biotinidase deficiency
397514364	BTD	NM_000060.3(BTD):c.485 C>T (p.Ala162Val)	CCTGAGTTGTATGGYCATCAGGG	Biotinidase deficiency
372844636	BTD	NM_000060.3(BTD):c.631 C>T (p.Arg211Cys)	CCCTTGTTGACCGCTACYGTAAA, CCTTGTTGACCGCTACYGTAAAC	Biotinidase deficiency
128621209	BTK	NM_000061.2(BTK):c.1838 G>A (p.Gly613Asp)	CCCAAGRCCTACGTCTCTACAGG	X-linked agammaglobulinemia
128621194	BTK	NM_000061.2(BTK):c.862 C>T (p.Arg288Trp)	CCAAACACATGACTYGGAGTCAG	X-linked agammaglobulinemia
128621204	BTK	NM_000061.2(BTK):c.1684 C>T (p.Arg562Trp)	CCAAATTTCCAGTCYGGTGGTCC	X-linked agammaglobulinemia
128621193	BTK	NM_000061.2(BTK):c.763 C>T (p.Arg255Ter)	CCATGGTGGAGAGCAYGAGATA A	X-linked agammaglobulinemia
137852956	C10orf2	NM_021830.4(C10orf2):c.9 08G>A (p.Arg303Gln)	GGCRGATTGTATTCTGGTTGGGG, CGGCRGATTGTATTCTGGTTGGG, CCGGCRGATTGTATTCTGGTTGG	Autosomal dominant progressive external ophthalmoplegia with mitochondrial DNA deletions 3
80356544	C10orf2	NM_021830.4(C10orf2):c.1 370C>T (p.Thr457Ile)	CCCGGGTCATGCTGAYACAGTTT, CCGGGTCATGCTGAYACAGTTTG	Mitochondrial DNA-depletion syndrome 3, hepatocerebral, Mitochondrial DNA depletion syndrome 7 (hepatocerebral type)
587777698	C12orf57	NM_138425.3(C12orf57):c. 184C>T (p.Gln62Ter)	CCCGTGGCCACGCAGATCYAGCA , CCGTGGCCACGCAGATCYAGCAG	Temtamy syndrome
397514539	C12orf65	NM_152269.4(C12orf65):c. 394C>T (p.Arg132Ter)	CCTGTTCACAAAGAAAAAYGAG A	Spastic paraplegia 55, autosomal recessive
397514477	C19orf12	NM_001031726.3(C19orf1	CCCTCGAAGGCCCGCCAYGATGA	Neurodegeneration with brain iron accumulation 4

		2):c.32C>T (p.Thr11Met)	, CCTCGAAGGCCCGCCAYGATGAC	
587777653	C2CD3	NM_001286577.1(C2CD3): c.184C>T (p.Arg62Ter)	ACTCRGACAAGTACACAAGTGGG , CACTCRGACAAGTACACAAGTGG	Joubert syndrome, Orofaciodigital syndrome xiv
121909587	C5	NM_001735.2(C5):c.55C>T (p.Gln19Ter)	CCTGGGGAAAACCTGGGGAYAG G	Leiner disease
139675596	C5orf42	NM_023073.3(C5orf42):c.7 477C>T (p.Arg2493Ter)	TCTGGTCRAAAAGTCACATTTGG	Joubert syndrome 17
121434552	CA4	NM_000717.3(CA4):c.206 G>A (p.Arg69His)	CTGGGACRCTTCTTCTTCTCTGG	Retinitis pigmentosa 17
104894559	CA4	NM_000717.3(CA4):c.40C >T (p.Arg14Trp)	CCTGGCCCTCTCCGCGGCGYGGC, CCCTCTCCGCGGCGYGGCCATCG	Retinitis pigmentosa 17
147623570	CA5A	NM_001739.1(CA5A):c.55 5G>A (p.Lys185=)	CCCGAGCAAGTGATTACYTTTAA, CCGAGCAAGTGATTACYTTTAAA	Carbonic anhydrase VA deficiency, hyperammonemia due to
121908215	CACNA1 A	NM_001127221.1(CACNA 1A):c.877G>A (p.Gly293Arg)	ACTGGGAARGGCCCAACAACGG G	Spinocerebellar ataxia 6, Episodic ataxia type 2
121908216	CACNA1 A	NM_001127221.1(CACNA 1A):c.4982G>A (p.Arg1661His)	TCCRCCTCTTCCGAGCTGCCCGG	Episodic ataxia type 2
121908236	CACNA1 A	NM_001127221.1(CACNA 1A):c.860G>A (p.Cys287Tyr)	AATRTCAGCCCTACTGGGAAGGG , AAATRTCAGCCCTACTGGGAAGG , GACCAAATRTCAGCCCTACTGGG	Episodic ataxia type 2
121908212	CACNA1 A	NM_001127221.1(CACNA 1A):c.1997C>T (p.Thr666Met)	CCCCTTTCAGATCCTGAYGGGCG, CCCTTTCAGATCCTGAYGGGCGA, CCTTTCAGATCCTGAYGGGCGAA	Familial hemiplegic migraine type 1
121909323	CACNA1 A	NM_001127221.1(CACNA 1A):c.3832C>T (p.Arg1278Ter)	CCCTTACAGGTGCTGYGATACTT, CCTTACAGGTGCTGYGATACTTT	Episodic ataxia type 2
121909324	CACNA1 A	NM_001127221.1(CACNA 1A):c.4636C>T (p.Arg1546Ter)	CCAAGCCGCTGACCYGACACATG	Episodic ataxia type 2
587776693	CACNA1 A	NM_001127221.1(CACNA 1A):c.3992+1G>A	CCCTTGCGAGGAGACTTAYGTGA , CCTTGCGAGGAGACTTAYGTGAA	Episodic ataxia type 2
80315385	CACNA1 C	NM_000719.6(CACNA1C): c.1204G>A (p.Gly402Ser)	TGGTTCTCRGTGTGCTTAGCGGG	Timothy syndrome, Congenital long QT syndrome

79891110	CACNA1C	NM_000719.6(CACNA1C):c.1216G>A (p.Gly406Arg)	TTAGCRGGTAAGCAGGACCAAGG	Timothy syndrome, Long QT syndrome, Congenital long QT syndrome, not provided
587782933	CACNA1C	NM_001167623.1(CACNA1C):c.1204G>A (p.Gly402Ser)	TGGTTCTCRGTGTGTTGAGCGGG	Paroxysmal familial ventricular fibrillation, not provided
122456133	CACNA1F	NM_005183.3(CACNA1F):c.1106G>A (p.Gly369Asp)	TTGRCGTCCTGAGTGGGTGAGGG, CTTGRCGTCCTGAGTGGGTGAGG	Congenital stationary night blindness, type 2A
122456135	CACNA1F	NM_005183.3(CACNA1F):c.2683C>T (p.Arg895Ter)	CCGCTGAGGACCCCATCYGAGCC	Congenital stationary night blindness, type 2A
80338777	CACNA1S	NM_000069.2(CACNA1S):c.1583G>A (p.Arg528His)	TCCRCTGCATCCGCCTCCTGAGG	Hypokalemic periodic paralysis 1
587777742	CACNB2	NM_201590.2(CACNB2):c.32C>T (p.Thr11Ile)	CCTTATAGCTCCTCAAAYTAAAT	Brugada syndrome 4
121917812	CACNB2	NM_201590.2(CACNB2):c.1442C>T (p.Ser481Leu)	CCGCTCTTCCTCCTYAGCCCCAC	Brugada syndrome 4
267606699	CANT1	NM_001159772.1(CANT1):c.899G>A (p.Arg300His)	CCTGCCGCRCCGCGCCAGCCAGG	Desbuquois syndrome
587776951	CANT1	NM_001159772.1(CANT1):c.-286+1G>A	CCGCGGGCGCAGTCACTCAYCCG	Desbuquois syndrome
377546036	CANT1	NM_001159772.1(CANT1):c.676G>A (p.Val226Met)	CCTTGCCCAGGCCGCCCAYGTAC	Desbuquois syndrome
141656719	CAPN3	NM_000070.2(CAPN3):c.1468C>T (p.Arg490Trp)	CCTGATGCAGAAGAACCGGYGGA	Limb-girdle muscular dystrophy, type 2A, not provided
121434546	CAPN3	NM_000070.2(CAPN3):c.257C>T (p.Ser86Phe)	CCCACCGGATGAGACCTYTCTCT, CCACCGGATGAGACCTYTCTCTT	Limb-girdle muscular dystrophy, type 2A
587777763	CARD14	NM_024110.4(CARD14):c.349+5G>A	GGTGARAGCTCCGACTTTGACGG	Psoriasis susceptibility 2
398122362	CARD9	NM_052813.4(CARD9):c.214G>A (p.Gly72Ser)	GACCRGCCACAAGGGCTACGTGG	Candidiasis, familial, 2
557671802	CARS2	NM_024537.3(CARS2):c.752C>T (p.Pro251Leu)	CCCRGCCTCCCGGGTCCCCAGGG, GCCCRGCCTCCCGGGTCCCCAGG	Alpers encephalopathy
794727270	CASK	NM_003688.3(CASK):c.79C>T (p.Arg27Ter)	CCCTTCAGTGTTGTAYGACGATG, CCTTCAGTGTTGTAYGACGATGT	FG syndrome 4, Mental retardation and microcephaly with pontine and cerebellar hypoplasia
587783361	CASK	NM_003688.3(CASK):c.2074C>T (p.Gln692Ter)	CCATGGAGAAGACCAAAYAGGAG	Mental retardation and microcephaly with pontine and cerebellar hypoplasia
587783364	CASK	NM_003688.3(CASK):c.2470C>T (p.Gln824Ter)	CCGGAAGATCCACGAGYAGGGGC	Mental retardation and microcephaly with pontine and cerebellar hypoplasia
587783371	CASK	NM_003688.3(CASK):c.880C>T (p.Gln294Ter)	CCAGAAACAGTAGAGYAGCTGAG	Mental retardation and microcephaly with pontine and cerebellar hypoplasia

28936699	CASP10	NM_032977.3(CASP10):c.1241C>T (p.Ala414Val)	CCTTCCGTATCCATCGAAGYAGA, CCGTATCCATCGAAGYAGATGCT	Malignant lymphoma, non-Hodgkin
104893700	CASR	NM_000388.3(CASR):c.2009G>A (p.Gly670Glu)	TCATCGRGGAGCCCCAGGACTGG	Hyperparathyroidism, neonatal severe
104893712	CASR	NM_000388.3(CASR):c.1810G>A (p.Glu604Lys)	GATCRAGTTTCTGTCGTGGACGG, AGGAGATCRAGTTTCTGTCGTGG	Hypocalcemia, autosomal dominant 1
104893719	CASR	NM_000388.3(CASR):c.1657G>A (p.Gly553Arg)	CAGGAAARGGATCATTGAGGGGG, CCAGGAAARGGATCATTGAGGGG	Hypocalciuric hypercalcemia, familial, type 1
121909264	CASR	NM_000388.3(CASR):c.428G>A (p.Gly143Glu)	GTGGTGGRAGCAACTGGCTCAGG	Hypocalciuric hypercalcemia, familial, type 1
121909266	CASR	NM_000388.3(CASR):c.196C>T (p.Arg66Cys)	CCAGGTATAATTTCYGTGGGTTT	Hypocalciuric hypercalcemia, familial, type 1
267606708	CBL	NM_005188.3(CBL):c.1259G>A (p.Arg420Gln)	TTCTGCCRATGTGAAATTAAAGG	Noonan syndrome-like disorder with or without juvenile myelomonocytic leukemia
28934891	CBS	NM_000071.2(CBS):c.1330G>A (p.Asp444Asn)	CTTCRACCAGGCGCCCGTGGTGG, GGGCTTCRACCAGGCGCCCGTGG	Homocystinuria due to CBS deficiency, Homocystinuria, pyridoxine-responsive, not provided
121908255	CBX2	NM_005189.2(CBX2):c.293C>T (p.Pro98Leu)	CCTTTATCTTTCCAGGAACYCGA	46,XY sex reversal, type 5
377177061	CC2D2A	NM_001080522.2(CC2D2A):c.394C>T (p.Arg132Ter)	CCTCGGCCCAGACGCTTAYGAAG	Meckel-Gruber syndrome
201133219	CCDC114	NM_144577.3(CCDC114):c.1391+5G>A	CCCTCCTGCCCTGCGCYTCACAG, CCTCCTGCCCTGCGCYTCACAGA	Ciliary dyskinesia, primary, 20
374909386	CCDC40	NM_017950.3(CCDC40):c.3354C>A (p.Tyr1118Ter)	CCGCGTGCGGGACGAGTAHCCCC	Ciliary dyskinesia, primary, 15
387907092	CCDC40	NM_017950.3(CCDC40):c.1951C>T (p.Gln651Ter)	CCACCAAATACTTCAACYAGCTC, CCAAATACTTCAACYAGCTCATC	Ciliary dyskinesia, primary, 15
587782989	CCDC88C	NM_001080414.3(CCDC88C):c.1391G>A (p.Arg464His)	GTCCAGCCRCATCCTGAAGCTGG	Spinocerebellar ataxia 40
387907320	CCDC88C	NM_001080414.3(CCDC88C):c.5058+1G>A	CCATRTGAGTGATCCGGACACGG	Hydrocephalus
137852841	CCM2	NM_001029835.2(CCM2):c.382C>T (p.Gln128Ter)	CCGGGACACTTGACTYAGGAGCA	Cerebral cavernous malformations 2
587777929	CCT7	NM_001166284.1(CCT7):c.1313C>T (p.Ser438Leu)	CCATCAAGAACCCCCGCTYGACT	Myocardial infarction 1
74315290	CD247	NM_198053.2(CD247):c.208C>T (p.Gln70Ter)	CCAGCAGGGCCAGAACYAGCTCT	Immunodeficiency due to defect in cd3-zeta

730880296	CD3D	NM_000732.4(CD3D):c.274+5G>A	TACRTGCTTCCTGAACCCTTTGG	Immunodeficiency 19
193922136	CD40LG	NM_000074.2(CD40LG):c.761C>T (p.Thr254Met)	CCATGGCACTGGCTTCAYGTCCT	Immunodeficiency with hyper IgM type 1
587776775	CD81	NM_004356.3(CD81):c.561+1G>A	CAAGRTGCGCGAGGCCGGTGGGG, TCAAGRTGCGCGAGGCCGGTGGG, TTCAAGRTGCGCGAGGCCGGTGG	
121918660	CD8A	NM_001768.6(CD8A):c.331G>A (p.Gly111Ser)	CGAGRGCTACTATTTCTGCTCGG	Cd8 deficiency, familial
113313967	CDAN1	NM_138477.2(CDAN1):c.1860+5G>A	CCCTTGTTCTGTTTTYGGACCTG, CCTTGTTCTGTTTTYGGACCTGC	Congenital dyserythropoietic anemia, type I
120074167	CDAN1	NM_138477.2(CDAN1):c.2015C>T (p.P672L)	CCCTCCCAGGTCCCTCYGGTCCT, CCTCCCAGGTCCCTCYGGTCCTG	Congenital dyserythropoietic anemia, type I
80338696	CDAN1	NM_138477.2(CDAN1):c.2140C>T (p.Arg714Trp)	CCCTTGCTGGAATATTACYGGGA, CCTTGCTGGAATATTACYGGGAC	Congenital dyserythropoietic anemia, type I
80338697	CDAN1	NM_138477.2(CDAN1):c.3124C>T (p.Arg1042Trp)	CCTTGGCCGTGGGGCCAYGGGAC	Congenital dyserythropoietic anemia, type I
121434263	CDC73	NM_024529.4(CDC73):c.128G>A (p.Trp43Ter)	GTTTRGGGGTAAGTCCGGCATGG	Parathyroid carcinoma
587776558	CDC73	NM_024529.4(CDC73):c.131+1G>A	GTTTGGGGRTAAGTCCGGCATGG	Hyperparathyroidism 1
587776559	CDC73	NM_024529.4(CDC73):c.238-1G>A	TTARACTGAAAATATTCCTGTGG	Hyperparathyroidism 2
786203576	CDH1	NM_004360.3(CDH1):c.60G>A (p.Trp20Ter)	CTCTTGRCTCTGCCAGGAGCCGG	Hereditary cancer-predisposing syndrome
121964875	CDH1	NM_004360.3(CDH1):c.59G>A (p.Trp20Ter)	CTCTTRGCTCTGCCAGGAGCCGG	Hereditary diffuse gastric cancer
121964877	CDH1	NM_004360.3(CDH1):c.1792C>T (p.Arg598Ter)	CCCCCATACCAGAACCTYGAACT, CCCCATACCAGAACCTYGAACTA, CCCATACCAGAACCTYGAACTAT, CCATACCAGAACCTYGAACTATA	Hereditary diffuse gastric cancer
587782750	CDH1	NM_004360.3(CDH1):c.1921C>T (p.Gln641Ter)	CCAACTGGACCATTYAGTACAAC	Hereditary cancer-predisposing syndrome
121434539	CDH15	NM_004933.2(CDH15):c.178C>T (p.Arg60Cys)	CCGAGAACCACAAGYGTCTCCCC	Mental retardation, autosomal dominant 3
397517353	CDH23	NM_022124.5(CDH23):c.7	CTGRGGCACCACCATGCTCCTGG	Usher syndrome, type 1D

		776G>A (p.Trp2592Ter)		
367928692	CDH23	NM_022124.5(CDH23):c.6050-9G>A	GCGGCACCRGGTGCCAGGTGTGG	Usher syndrome, type 1D
727502931	CDH23	NM_022124.5(CDH23):c.7362+5G>A	GTGARCAGTGATGGAGGGCCTGG	Usher syndrome, type 1D
121908354	CDH23	NM_022124.5(CDH23):c.719C>T (p.Pro240Leu)	CCCATCTTCATCAACCTGCYTTA, CCATCTTCATCAACCTGCYTTAC	Deafness, autosomal recessive 12
727503845	CDK16	NM_033018.3(CDK16):c.1258C>T (p.Arg420Ter)	CCTTTTGAGCCACGCACCCYGGT	not provided
587783392	CDK5RAP2	NM_018249.5(CDK5RAP2):c.5227C>T (p.Gln1743Ter)	CCTGCTCAAACAGATCAGCYAGG	Primary autosomal recessive microcephaly 3
587783086	CDKL5	NM_003159.2(CDKL5):c.577G>A (p.Asp193Asn)	TCCGTGRACATGTGGTCGGTGGG, GTCCGTGRACATGTGGTCGGTGG	not provided
267608653	CDKL5	NM_003159.2(CDKL5):c.2152G>A (p.Val718Met)	TACAGARGTAAGCCCACCCCCGG	Early infantile epileptic encephalopathy 2, not provided
122460158	CDKL5	NM_003159.2(CDKL5):c.2500C>T (p.Gln834Ter)	CCTTTCTTTCAGAGCYAGCCATT	Early infantile epileptic encephalopathy 2, Atypical Rett syndrome
61749704	CDKL5	NM_003159.2(CDKL5):c.539C>T (p.Pro180Leu)	CCAGATGGTATCGGTCCCYAGAA	Early infantile epileptic encephalopathy 2
267606713	CDKL5	NM_003159.2(CDKL5):c.863C>T (p.Thr288Ile)	CCAGCTGACAGATACTTGAYAGA	Early infantile epileptic encephalopathy 2
587783089	CDKL5	NM_003159.2(CDKL5):c.700C>T (p.Gln234Ter)	CCACTTCCATCTGAGYAGATGAA	not provided
587783158	CDKL5	NM_003159.2(CDKL5):c.2596C>T (p.Gln866Ter)	CCAGCCCTTAACAGCTCAAYAAA, CCCTTAACAGCTCAAYAAACCAA, CCTTAACAGCTCAAYAAACCAAA	Early infantile epileptic encephalopathy 2, not provided
267608643	CDKL5	NM_003159.2(CDKL5):c.1648C>T (p.Arg550Ter)	CCCTTCTGGAAGAAATAACYGAA, CCTTCTGGAAGAAATAACYGAAA	Early infantile epileptic encephalopathy 2, Atypical Rett syndrome, not provided
267608659	CDKL5	NM_003159.2(CDKL5):c.2413C>T (p.Gln805Ter)	CCCTGATCTTCTGACGTTGYAGA, CCTGATCTTCTGACGTTGYAGAA	Early infantile epileptic encephalopathy 2
267608663	CDKL5	NM_003159.2(CDKL5):c.2593C>T (p.Gln865Ter)	CCAGCCCTTAACAGCTYAACAAA	Early infantile epileptic encephalopathy 2
121917832	CDKN1B	NM_004064.4(CDKN1B):c.227G>A (p.Trp76Ter)	GAGTRGCAAGAGGTGGAGAAGGG, CGAGTRGCAAGAGGTGGAGAAG	Multiple endocrine neoplasia, type 4

			G	
387907225	CDKN1C	NM_000076.2(CDKN1C):c.820G>A (p.Asp274Asn)	GATCTCCRGTGAGCCCCGCACGG	Intrauterine growth retardation, metaphyseal dysplasia, adrenal hypoplasia congenita, and genital anomalies
387906918	CDT1	NM_030928.3(CDT1):c.196G>A (p.Ala66Thr)	CACCGRCCCGCAGGAGACTGCGG	Meier-Gorlin syndrome 4
145646425	CEP164	NM_014956.4(CEP164):c.1726C>T (p.Arg576Ter)	CCCCACCCATGGTAGGYGATCCA, CCCACCCATGGTAGGYGATCCAC, CCACCCATGGTAGGYGATCCACA	Nephronophthisis 15
387907310	CEP164	NM_014956.4(CEP164):c.277C>T (p.Arg93Trp)	CCATGTGACGAACACTATYGGAG	Nephronophthisis 15
387907311	CEP164	NM_014956.4(CEP164):c.1573C>T (p.Gln525Ter)	CCTGCAGCTGTCCCTCYAGAGGT	Nephronophthisis 15
371812716	CEP41	NM_018718.2(CEP41):c.1078C>T (p.Arg360Cys)	GGAGCRGGGGTTTGAGTGGCTGG	Joubert syndrome 9/15, digenic
375801610	CFAP53	NM_145020.4(CFAP53):c.121C>T (p.Arg41Ter)	TGCGTCRGATTCTTTCTAGATGG	Heterotaxy, visceral, 6, autosomal
398123065	CFB	NM_001710.5(CFB):c.766C>T (p.Gln256Ter)	CCTGGCACCCAGGGGAAYAACAG	Complement factor B deficiency
104893611	CFC1	NM_032545.3(CFC1):c.334C>T (p.Arg112Cys)	CCCGGCCCACTTCACCGGCYGCT, CCGGCCCACTTCACCGGCYGCTA, CCCACTTCACCGGCYGCTACTGC	Heterotaxy, visceral, 2, autosomal
121913053	CFH	NM_000186.3(CFH):c.2876G>A (p.Cys959Tyr)	TACAAATRTTTTGAAGGTTTTGG	Factor H deficiency
121964916	CFI	NM_000204.3(CFI):c.728G>A (p.Gly243Asp)	TGTGATGRTATCAATGATTGTGG	Afibrinogenemia
132630258	CFP	NM_002621.2(CFP):c.481C>T (p.Arg161Ter)	CCCGGACCCGCAGGYGAGCCTGT	Properdin deficiency, X-linked
672601317	CFTR	NM_000492.3(CFTR):c.830G>A (p.Trp277Ter)	ATACTGCTRGGAAGAAGCAATGG	Cystic fibrosis
121908753	CFTR	NM_000492.3(CFTR):c.1055G>A (p.Arg352Gln)	GGTCACTCRGCAATTTCCCTGGG	Cystic fibrosis
121909010	CFTR	NM_000492.3(CFTR):c.3947G>A (p.Trp1316Ter)	AATATRGAAAGTTGCAGATGAGG	Cystic fibrosis
397508200	CFTR	NM_000492.3(CFTR):c.1393-1G>A	TTTCCARACTTCACTTCTAATGG	Cystic fibrosis
387906369	CFTR	NM_000492.3(CFTR):c.3718-1G>A	ACCTTATARGTGGGCCTCTTGGG	Cystic fibrosis

397508256	CFTR	NM_000492.3(CFTR):c.166G>A (p.Glu56Lys)	CAGARAATGGGATAGAGAGCTGG	Cystic fibrosis
397508279	CFTR	NM_000492.3(CFTR):c.170G>A (p.Trp57Ter)	CAGAGAATRGGATAGAGAGCTGG	Cystic fibrosis
77409459	CFTR	NM_000492.3(CFTR):c.1013C>T (p.Thr338Ile)	CCTCCGGAAAATATTCAYCACCA, CCGGAAAATATTCAYCACCATCT	Cystic fibrosis
121908760	CFTR	NM_000492.3(CFTR):c.2125C>T (p.Arg709Ter)	CCAATCAACTCTATAYGAAAATT	Cystic fibrosis
121908802	CFTR	NM_000492.3(CFTR):c.595C>T (p.His199Tyr)	CCAGGGACTTGCATTGGCAYATT	Cystic fibrosis
121908810	CFTR	NM_000492.3(CFTR):c.2290C>T (p.Arg764Ter)	CCCCACGCTTCAGGCAYGAAGGA, CCCACGCTTCAGGCAYGAAGGAG, CCACGCTTCAGGCAYGAAGGAGG	Cystic fibrosis
374946172	CFTR	NM_000492.3(CFTR):c.2353C>T (p.Arg785Ter)	CCAAGGTCAGAACATTCACYGAA	Cystic fibrosis
121912816	CHAT	NM_020549.4(CHAT):c.1321G>A (p.Glu441Lys)	TGCRAACACTCCCCATTCGATGG	Familial infantile myasthenia
121912819	CHAT	NM_020549.4(CHAT):c.1679G>A (p.Arg560His)	TCCATCCRCCGATTCCAGGAGGG, GTCCATCCRCCGATTCCAGGAGG	Familial infantile myasthenia
794727516	CHAT	NM_020549.4(CHAT):c.418C>T (p.Gln140Ter)	CCCGTGCCCCCGCTGYAGCAGAC, CCGTGCCCCCGCTGYAGCAGACC	Familial infantile myasthenia
121912821	CHAT	NM_020549.4(CHAT):c.1493C>T (p.Ser498Leu)	CCACTTAGCCTCCTYGGCAGAAA	Familial infantile myasthenia
398123000	CHD2	NM_001271.3(CHD2):c.1396C>T (p.Arg466Ter)	CCCTGAAGCAGAGACCAYGATTT, CCTGAAGCAGAGACCAYGATTTG	Epileptic encephalopathy, childhood-onset
727503863	CHD7	NM_017780.3(CHD7):c.2933G>A (p.Trp978Ter)	TAAACTRGCTACTTTTCAATTGG	CHARGE association
121434343	CHD7	NM_017780.3(CHD7):c.6322G>A (p.Gly2108Arg)	TAAACACRGGGTCAGTCGGACGG	CHARGE association
587783429	CHD7	NM_017780.3(CHD7):c.1480C>T (p.Arg494Ter)	CCCACAAGCAATCCAGGAAYGAC, CCACAAGCAATCCAGGAAYGACT	CHARGE association
727503861	CHD7	NM_017780.3(CHD7):c.1369C>T (p.Gln457Ter)	CCCCAGAAACATGCAGYAGTCTC,	not provided

			CCCAGAAACATGCAGYAGTCTCG, CCAGAAACATGCAGYAGTCTCGT	
587783440	CHD7	NM_017780.3(CHD7):c.4318C>T (p.Gln1440Ter)	CCTGGATAAAGCTGTGCTAYAGT	CHARGE association
587783458	CHD7	NM_017780.3(CHD7):c.7957C>T (p.Arg2653Ter)	CCTGTTGTCAATAAAYGAAATGG	CHARGE association
786203889	CHEK2	NM_007194.3(CHEK2):c.278G>A (p.Trp93Ter)	TGCCCCCTRGGCTCGATTATGGG	Hereditary cancer-predisposing syndrome
587781269	CHEK2	NM_007194.3(CHEK2):c.283C>T (p.Arg95Ter)	CCCCTGCCCCCTGGGCTYGATTA, CCCTGCCCCCTGGGCTYGATTAT, CCTGCCCCCTGGGCTYGATTATG	Hereditary cancer-predisposing syndrome
17883862	CHEK2	NM_007194.3(CHEK2):c.254C>T (p.Pro85Leu)	CCTGAGGACCAAGAACYTGAGGA	Familial cancer of breast, Hereditary cancer-predisposing syndrome, Osteosarcoma, not specified
63750355	CHMP2B	NM_014043.3(CHMP2B):c.493C>T (p.Gln165Ter)	CCAGGATATTGTGAATYAAGTTC	Frontotemporal Dementia, Chromosome 3-Linked, not provided
121912796	CHN1	NM_001822.5(CHN1):c.682G>A (p.Gly228Ser)	ATGTGGRGTCTCATTGCTCAGGG, TATGTGGRGTCTCATTGCTCAGG	Duane syndrome type 2
121912798	CHN1	NM_001822.5(CHN1):c.937G>A (p.Glu313Lys)	CCTAATTRAAGATGTCAAGATGG	Duane syndrome type 2
387906599	CHN1	NM_001822.5(CHN1):c.422C>T (p.Pro141Leu)	CCAAGATGACGATAAACCYAATT	Duane syndrome type 2
281865066	CHRNA4	NM_000744.6(CHRNA4):c.878C>T (p.Thr293Ile)	CCTGCTGCTCATCAYCGAGATCA	Epilepsy, nocturnal frontal lobe, type 1
137852810	CHRNB1	NM_000747.2(CHRNB1):c.865G>A (p.Val289Met)	TACTRTGTTCCTGCTGCTGCTGG	MYASTHENIC SYNDROME, CONGENITAL, 2A, SLOW-CHANNEL
121912672	CHRNG	NM_005199.4(CHRNG):c.136C>T (p.Arg46Ter)	CCTGCGGCCCGCGGAAYGAGACT	Multiple pterygium syndrome Escobar type
267606725	CHRNG	NM_005199.4(CHRNG):c.13C>T (p.Gln5Ter)	CCATGCATGGGGGCYAGGGGCCG	Multiple pterygium syndrome Escobar type
267606734	CHST3	NM_004273.4(CHST3):c.1114G>A (p.Glu372Lys)	GCTACRAGGACGTGGCACGCGGG, CGCTACRAGGACGTGGCACGCGG	Spondyloepiphyseal dysplasia with congenital joint dislocations
80356700	CLCN1	NM_000083.2(CLCN1):c.689G>A (p.Gly230Glu)	CGTGGRGAAAGAGGTAGGCCTGG	Myotonia congenita, Congenital myotonia, autosomal dominant form
80356702	CLCN1	NM_000083.2(CLCN1):c.950G>A (p.Arg317Gln)	TGTTTCRAGTGCTGGCAGTGTGG	Myotonia congenita, Congenital myotonia, autosomal recessive form, Congenital myotonia, autosomal dominant form
80356693	CLCN1	NM_000083.2(CLCN1):c.1	CCAGTTCTGGATGTYCATCGTGG	Myotonia congenita

		412C>T (p.Ser471Phe)		
80356706	CLCN1	NM_000083.2(CLCN1):c.2795C>T (p.Pro932Leu)	CCCCAGAGACCCCTGTGCYATCT, CCCAGAGACCCCTGTGCYATCTC, CCAGAGACCCCTGTGCYATCTCC	Myotonia congenita, Congenital myotonia, autosomal recessive form
80356694	CLCN1	NM_000083.2(CLCN1):c.1439C>T (p.Pro480Leu)	CCACCACTATGCCCATACYCTGC, CCACTATGCCCATACYCTGCGGA	Myotonia congenita, Congenital myotonia, autosomal dominant form
201330912	CLCN2	NM_004366.5(CLCN2):c.1709G>A (p.Trp570Ter)	CCGTACTGGTGGCGGCCCYAGCC	Leukoencephalopathy with ataxia
151340625	CLCN5	NM_001127899.3(CLCN5):c.1727G>A (p.Gly576Glu)	GTTGRGGCTGCAGCCTGCTTAGG	
121434433	CLCN7	NM_001287.5(CLCN7):c.2285G>A (p.Arg762Gln)	TGTTCCRGGCCCTGGGCCTGCGG	Osteopetrosis autosomal recessive 4
121434432	CLCN7	NM_001287.5(CLCN7):c.1663C>T (p.Gln555Ter)	CCTGATGGGAGCTGCTGCCYAGC	Osteopetrosis autosomal recessive 4
121434435	CLCN7	NM_001287.5(CLCN7):c.2299C>T (p.Arg767Trp)	CCGGGCCCTGGGCCTGYGGCACC	Osteopetrosis autosomal dominant type 2, Osteopetrosis autosomal recessive 4
121909132	CLCNKB	NM_000085.4(CLCNKB):c.610G>A (p.Ala204Thr)	GGCARCGGCGGCAGTGGGCGTGG	Bartter syndrome type 3
121909136	CLCNKB	NM_000085.4(CLCNKB):c.1830G>A (p.Trp610Ter)	CTGRGCTCCTGGACACCAGGTGG, TTCCTGRGCTCCTGGACACCAGG	Bartter syndrome, type 3, with hypocalciuria
121909131	CLCNKB	NM_000085.4(CLCNKB):c.371C>T (p.Pro124Leu)	CCAAGGTTCTGGAATCCYGGAGG	Bartter syndrome type 3
104893721	CLDN16	NM_006580.3(CLDN16):c.715G>A (p.Gly239Arg)	CTCRGAATGGCTGGGTCTCTGGG, GCTCRGAATGGCTGGGTCTCTGG	Primary hypomagnesemia
104893723	CLDN16	NM_006580.3(CLDN16):c.593G>A (p.Gly198Asp)	TCAGRTACCCCAGGAATCATTGG	Primary hypomagnesemia
104893727	CLDN16	NM_006580.3(CLDN16):c.698G>A (p.Gly233Asp)	TTTGRTTGGTCCTGTTGGCTCGG	Primary hypomagnesemia
796052335	CLN3	NM_001042432.1(CLN3):c.949C>T (p.Gln317Ter)	CCCTGAGTCACGCTYAGCAATAC	not provided
104894484	CLN6	NM_017882.2(CLN6):c.368G>A (p.Gly123Asp)	CATGGRTGCCAGCATCCACCTGG	Ceroid lipofuscinosis neuronal 6
796052356	CLN6	NM_017882.2(CLN6):c.665+1G>A	CTGRTGAGTGGACATCAGCATGG	not provided
154774635	CLN6	NM_017882.2(CLN6):c.139C>T (p.Leu47Phe)	CCCTTCCACCTCGACYTCTGGTT, CCTTCCACCTCGACYTCTGGTTC	Adult neuronal ceroid lipofuscinosis, not provided
104893615	CNGA3	NM_001298.2(CNGA3):c.1669G>A (p.Gly557Arg)	ACATCAAGRGGAGCAAGTCGGGG	Achromatopsia 2, not specified
104893619	CNGA3	NM_001298.2(CNGA3):c.1	GGCCRTGGTGGCTGATGATGGGG	Achromatopsia 2

		585G>A (p.Val529Met)	, TGGCCRTGGTGGCTGATGATGGG, CTGGCCRTGGTGGCTGATGATGG	
104893613	CNGA3	NM_001298.2(CNGA3):c.847C>T (p.Arg283Trp)	CCGCCTACTGAAGTTTTCCYGGC, CCTACTGAAGTTTTCCYGGCTCT	Achromatopsia 2
104893620	CNGA3	NM_001298.2(CNGA3):c.829C>T (p.Arg277Cys)	CCCAGAAGTGAGGTTCAACYGCC , CCAGAAGTGAGGTTCAACYGCCT	Achromatopsia 2
104893621	CNGA3	NM_001298.2(CNGA3):c.1306C>T (p.Arg436Trp)	CCAAGGACTTGGAGACGYGGGTT	Achromatopsia 2
372504780	CNGB1	NM_001297.4(CNGB1):c.952C>T (p.Gln318Ter)	TCCTRGTGGGCATCCTCCCAGGG, ATCCTRGTGGGCATCCTCCCAGG	Retinitis pigmentosa 45, not provided
786205909	CNNM2	NM_017649.4(CNNM2):c.364G>A (p.Glu122Lys)	CCTTCACCRAGCACGAGCGGCGG	HYPOMAGNESEMIA, SEIZURES, AND MENTAL RETARDATION
786205910	CNNM2	NM_017649.4(CNNM2):c.1069G>A (p.Glu357Lys)	CTTCGGARAGATCGTGCCCCAGG	HYPOMAGNESEMIA, SEIZURES, AND MENTAL RETARDATION
80100937	CNNM4	NM_020184.3(CNNM4):c.1690C>T (p.Gln564Ter)	CCTCCTACAGAGGTCTCTYAGTT, CCTACAGAGGTCTCTYAGTTTAG	Cone-rod dystrophy amelogenesis imperfecta
398124268	CNTNAP2	NM_014141.5(CNTNAP2):c.2153G>A (p.Trp718Ter)	TACTRGGGAGGCTCTGGGCCTGG	not provided
587777136	COASY	NM_025233.6(COASY):c.175C>T (p.Gln59Ter)	CCCCAGTCCAGCCCCGTGYAGGC , CCCAGTCCAGCCCCGTGYAGGCC , CCAGTCCAGCCCCGTGYAGGCCA	Neurodegeneration with brain iron accumulation 6
267606740	COG4	NM_015386.2(COG4):c.2197C>T (p.Arg733Trp)	CCGAGACAAGTTTGCCYGGCTCT	Congenital disorder of glycosylation type 2J
121912946	COL11A2	NM_080680.2(COL11A2):c.4322G>A (p.Gly1441Glu)	CCCTGRGCAGAAGGGTGAGATGG	Weissenbacher-Zweymuller syndrome
121912947	COL11A2	NM_080680.2(COL11A2):c.3100C>T (p.Arg1034Cys)	CCCATTGGTCCGCCAGGGYGCCC, CCATTGGTCCGCCAGGGYGCCCA	Deafness, autosomal dominant 13
121912951	COL11A2	NM_080680.2(COL11A2):c.3991C>T (p.Arg1331Ter)	CCTGGTTCCGAGGGGYGACAAGG	
200487396	COL12A1	NM_004370.5(COL12A1):c.5893C>T (p.Arg1965Cys)	ACAACGCRATATTGCAGCACAGG	BETHLEM MYOPATHY 2
796052094	COL12A1	NM_004370.5(COL12A1):c.8357G>A (p.Gly2786Asp)	CCAGRCCCCCAGGGTCCTCCAGG	BETHLEM MYOPATHY 2
121912773	COL17A1	NM_000494.3(COL17A1):c.1898G>A (p.Gly633Asp)	CGTGRTGAGGCAGGGCCTCCTGG	Adult junctional epidermolysis bullosa

121912769	COL17A1	NM_000494.3(COL17A1):c.3676C>T (p.Arg1226Ter)	CCTGGTCCCCCAGGGCCTYGAGG	Adult junctional epidermolysis bullosa, Epidermolysis bullosa, junctional, localisata variant
72648320	COL1A1	NM_000088.3(COL1A1):c.1200+1G>A	GCCAATRTAAGTATCCTGCCAGG	Osteogenesis imperfecta
72648356	COL1A1	NM_000088.3(COL1A1):c.1598G>A (p.Gly533Asp)	AGCTGRTCTGCCTGGTGCCAAGG	Osteogenesis imperfecta, recessive perinatal lethal
72651646	COL1A1	NM_000088.3(COL1A1):c.2156G>A (p.Gly719Asp)	CCCGRTAGCCAGGGCGCCCCTGG	Osteogenesis imperfecta, recessive perinatal lethal
72651651	COL1A1	NM_000088.3(COL1A1):c.2210G>A (p.Gly737Asp)	GCTGRTCTTCCAGGGCCTAAGGG, AGCTGRTCTTCCAGGGCCTAAGG	Osteogenesis imperfecta, recessive perinatal lethal
72653131	COL1A1	NM_000088.3(COL1A1):c.2515G>A (p.Gly839Ser)	GCTRGTCCCCCTGGCCCTGCCGG	Osteogenesis imperfecta type III
72653136	COL1A1	NM_000088.3(COL1A1):c.2533G>A (p.Gly845Arg)	GCCRGACCCGCTGGACCCCCTGG	Osteogenesis imperfecta, recessive perinatal lethal
72653137	COL1A1	NM_000088.3(COL1A1):c.2552G>A (p.Gly851Asp)	CCCCCTGRCCCCATTGTGAGTGG	Osteogenesis imperfecta, recessive perinatal lethal
72653169	COL1A1	NM_000088.3(COL1A1):c.3028G>A (p.Gly1010Ser)	CCCTRGTGAATCTGGACGTGAGG	Osteogenesis imperfecta with normal sclerae, dominant form
72653172	COL1A1	NM_000088.3(COL1A1):c.3073G>A (p.Gly1025Arg)	CCTRGACGAGACGGTTCTCCTGG	Osteogenesis imperfecta, recessive perinatal lethal
72653178	COL1A1	NM_000088.3(COL1A1):c.3118G>A (p.Gly1040Ser)	ACCRGCCCCGCTGGACCCCCTGG	Osteogenesis imperfecta type III
72654797	COL1A1	NM_000088.3(COL1A1):c.3182G>A (p.Gly1061Asp)	GCTGRCAAGAGTGGTGATCGTGG	Osteogenesis imperfecta, recessive perinatal lethal
72654802	COL1A1	NM_000088.3(COL1A1):c.3235G>A (p.Gly1079Ser)	GTCRGCCCTGTTGGCGCCCGTGG	Osteogenesis imperfecta type I
72656306	COL1A1	NM_000088.3(COL1A1):c.3271G>A (p.Gly1091Ser)	CAARGCCCCCGTGGTGACAAGGG, CCAARGCCCCCGTGGTGACAAGG	Osteogenesis imperfecta, recessive perinatal lethal
72656330	COL1A1	NM_000088.3(COL1A1):c.3541G>A (p.Gly1181Ser)	CCCRGCCCTCCTGGACCTCCTGG	Osteogenesis imperfecta, recessive perinatal lethal
66523073	COL1A1	NM_000088.3(COL1A1):c.3064G>A (p.Gly1022Ser)	GAARGTTCCCCTGGACGAGACGG	Osteogenesis imperfecta type III
72645320	COL1A1	NM_000088.3(COL1A1):c.761G>A (p.Gly254Glu)	CGAGRATTGCCCGGAACAGCTGG	Osteogenesis imperfecta type III
72645333	COL1A1	NM_000088.3(COL1A1):c.824G>A (p.Gly275Asp)	GATGRTGCCAAGGGAGATGCTGG	Osteogenesis imperfecta, recessive perinatal lethal
72645357	COL1A1	NM_000088.3(COL1A1):c.994G>A (p.Gly332Arg)	TGCCRGGCCCCCTGTGAGTGTGG	Osteogenesis imperfecta, Osteogenesis imperfecta type III

72653170	COL1A1	NM_000088.3(COL1A1):c.3040C>T (p.Arg1014Cys)	CCCCCTGGTGAATCTGGAYGTGA, CCCCTGGTGAATCTGGAYGTGAG , CCCTGGTGAATCTGGAYGTGAGG , CCTGGTGAATCTGGAYGTGAGGT	Infantile cortical hyperostosis
72653173	COL1A1	NM_000088.3(COL1A1):c.3076C>T (p.Arg1026Ter)	CCGAAGGTTCCCCTGGAYGAGAC	Osteogenesis imperfecta
72645347	COL1A1	NM_000088.3(COL1A1):c.934C>T (p.Arg312Cys)	CCTGCCTGGTGAGAGAGGTYGCC , CCTGGTGAGAGAGGTYGCCCTGG	Ehlers-Danlos syndrome, classic type
72658152	COL1A2	NM_000089.3(COL1A2):c.1981G>A (p.Gly661Ser)	CCTRGTCTCAGAGGTGAAATTGG	Osteoporosis
72658161	COL1A2	NM_000089.3(COL1A2):c.2099G>A (p.Gly700Asp)	GCTGRTCCTGCTGGTCCTGCTGG	Osteogenesis imperfecta type III, Osteogenesis imperfecta with normal sclerae, dominant form
72658176	COL1A2	NM_000089.3(COL1A2):c.2251G>A (p.Gly751Ser)	AACRGTGTTGTTGGTCCCACAGG	Osteogenesis imperfecta type III
72658200	COL1A2	NM_000089.3(COL1A2):c.2575G>A (p.Gly859Ser)	CCTRGCACTCCAGGTCCTCAGGG, TCCTRGCACTCCAGGTCCTCAGG	Osteogenesis imperfecta type III
72659338	COL1A2	NM_000089.3(COL1A2):c.3295G>A (p.Gly1099Arg)	CCTRGACCTCCAGGTGTAAGCGG	Osteogenesis imperfecta type III
121912900	COL1A2	NM_000089.3(COL1A2):c.2720G>A (p.Gly907Asp)	CGTGRTCCTCCTGGTGCTGTGGG, CCGTGRTCCTCCTGGTGCTGTGG	Osteogenesis imperfecta, recessive perinatal lethal
121912901	COL1A2	NM_000089.3(COL1A2):c.1640G>A (p.Gly547Asp)	CAGGRTCCCCCTGGTCCTCCAGG	Osteogenesis imperfecta, recessive perinatal lethal
121912902	COL1A2	NM_000089.3(COL1A2):c.2593G>A (p.Gly865Ser)	CAGRGTCTTCTTGGTGCTCCTGG	Osteogenesis imperfecta, recessive perinatal lethal
121912904	COL1A2	NM_000089.3(COL1A2):c.2414G>A (p.Gly805Asp)	TCTGRCCCTCCTGGTCCCCCTGG	Osteogenesis imperfecta, recessive perinatal lethal
121912909	COL1A2	NM_000089.3(COL1A2):c.1739G>A (p.Gly580Asp)	TTTGRTCTCCCTGGTCCTGCTGG	Osteogenesis imperfecta, recessive perinatal lethal
121912910	COL1A2	NM_000089.3(COL1A2):c.1504G>A (p.Gly502Ser)	TAGRGTGATCCTGGCAAAAACGG	Osteogenesis imperfecta, recessive perinatal lethal
267606741	COL1A2	NM_000089.3(COL1A2):c.1262G>A (p.Gly421Asp)	CCTGRTAGTCGTGGTGCAAGTGG	Osteogenesis imperfecta, recessive perinatal lethal
267606742	COL1A2	NM_000089.3(COL1A2):c.3269G>A (p.Gly1090Asp)	CAGGRCCCCCCTGGTCCCCCTGG	Osteogenesis imperfecta type III
72656387	COL1A2	NM_000089.3(COL1A2):c.838G>A (p.Gly280Ser)	GCCRGTCCCCGTGGTGAAGTGGG ,	Osteogenesis imperfecta

			CGCCRGTCCCCGTGGTGAAGTGG	
121912864	COL2A1	NM_001844.4(COL2A1):c.3220G>A (p.Gly1074Ser)	CCTRGCTCCCCTGGCCCCGCTGG	Hypochondrogenesis
121912867	COL2A1	NM_001844.4(COL2A1):c.2320G>A (p.Gly774Ser)	AAARGCCCTGAGGGAGCCCCTGG	Hypochondrogenesis
121912872	COL2A1	NM_001844.4(COL2A1):c.800G>A (p.Gly267Asp)	AAGGGRTCCGCCTGGTCCTCAGG	STICKLER SYNDROME, TYPE I, NONSYNDROMIC OCULAR
121912877	COL2A1	NM_001844.4(COL2A1):c.908G>A (p.Gly303Asp)	GGCGGRTGCTCCTGGTGTGAAGG	Stickler syndrome type 1, Kniest dysplasia
121912878	COL2A1	NM_001844.4(COL2A1):c.2905G>A (p.Gly969Ser)	GAARGTCCACCAGGTCCCCAGGG, CGAARGTCCACCAGGTCCCCAGG	Achondrogenesis type 2
121912888	COL2A1	NM_001844.4(COL2A1):c.1547G>A (p.Gly516Asp)	CGCGRTTTCCCAGGTCAAGATGG	Achondrogenesis type 2
121912891	COL2A1	NM_001844.4(COL2A1):c.3508G>A (p.Gly1170Ser)	GTCRGTCCCTCTGGCAAAGATGG	Coxa plana
121912894	COL2A1	NM_001844.4(COL2A1):c.952G>A (p.Gly318Arg)	GAACRGATCTCCGGGCCCAATGG	Rhegmatogenous retinal detachment, autosomal dominant
121912896	COL2A1	NM_001844.4(COL2A1):c.141G>A (p.Trp47Ter)	TGTGRAAGCCGGAGCCCTGCCGG	STICKLER SYNDROME, TYPE I, NONSYNDROMIC OCULAR
138498898	COL2A1	NM_001844.4(COL2A1):c.4148C>T (p.Thr1383Met)	CTTCCRTGGACAGCAGGCGTAGG	
121912868	COL2A1	NM_001844.4(COL2A1):c.3158G>A (p.Gly1053Glu)	CTGRAGTCAAGGTGAGTGTCTGG	Hypochondrogenesis
387906558	COL2A1	NM_001844.4(COL2A1):c.2149G>A (p.Gly717Ser)	CAGRGTCCCCGTGGCCTCCCCGG	
121912865	COL2A1	NM_001844.4(COL2A1):c.2155C>T (p.Arg719Cys)	CCAGGGCCTCCAGGGTCCCYGTG	Osteoarthritis with mild chondrodysplasia
121912882	COL2A1	NM_001844.4(COL2A1):c.2710C>T (p.Arg904Cys)	CCCTGGAGCTGCTGGCYGCGTTG, CCTGGAGCTGCTGGCYGCGTTGG	Epiphyseal dysplasia, multiple, with myopia and conductive deafness
113871730	COL3A1	NM_000090.3(COL3A1):c.926G>A (p.Gly309Glu)	GAGRACGGCCAGGACTTCCTGGG, CGAGRACGGCCAGGACTTCCTGG	Ehlers-Danlos syndrome, type 4
113485686	COL3A1	NM_000090.3(COL3A1):c.2356G>A (p.Gly786Arg)	CCCRGACTTCCAGGTATAGCTGG	Ehlers-Danlos syndrome, type 4, Ehlers-Danlos syndrome, type 4 variant
121912916	COL3A1	NM_000090.3(COL3A1):c.3041G>A (p.Gly1014Glu)	CAGGRAAACCCTGGATCAGATGG	Ehlers-Danlos syndrome, type 4
121912919	COL3A1	NM_000090.3(COL3A1):c.907G>A (p.Gly303Arg)	AGARGGGCTCCTGGTGAGCGAGG	Ehlers-Danlos syndrome, type 4

121912920	COL3A1	NM_000090.3(COL3A1):c.2410G>A (p.Gly804Ser)	ACTRGCCCTCCAGGACCTGCTGG	
121912921	COL3A1	NM_000090.3(COL3A1):c.1997G>A (p.Gly666Asp)	GCCGRTGCACCTGGAGCTCCAGG	Ehlers-Danlos syndrome, type 4
121912924	COL3A1	NM_000090.3(COL3A1):c.3302G>A (p.Gly1101Glu)	CGTGRAGCTGCTGGCATCAAAGG	Ehlers-Danlos syndrome, type 4
587779419	COL3A1	NM_000090.3(COL3A1):c.1033G>A (p.Gly345Arg)	CCCTRGATCCCCTGGTGCTAAGG	Ehlers-Danlos syndrome, type 4
587779432	COL3A1	NM_000090.3(COL3A1):c.2780G>A (p.Gly927Asp)	GCTGRCCAACCAGGAGAGAAGGG, TGCTGRCCAACCAGGAGAGAAGG	Ehlers-Danlos syndrome, type 4
587779434	COL3A1	NM_000090.3(COL3A1):c.2861G>A (p.Gly954Glu)	ACTGRAGCACGGGGTCTTGCAGG	Ehlers-Danlos syndrome, type 4
587779437	COL3A1	NM_000090.3(COL3A1):c.2140G>A (p.Gly714Arg)	CCTRGGCCACCTGGTGCTGCTGG	Ehlers-Danlos syndrome, type 4
587779438	COL3A1	NM_000090.3(COL3A1):c.2824G>A (p.Gly942Arg)	TAGRGAGCTCCAGGCCCACTTGG	Ehlers-Danlos syndrome, type 4
587779439	COL3A1	NM_000090.3(COL3A1):c.3301G>A (p.Gly1101Arg)	CGTRGAGCTGCTGGCATCAAAGG	Ehlers-Danlos syndrome, type 4
587779446	COL3A1	NM_000090.3(COL3A1):c.556G>A (p.Gly186Ser)	CCTRGTACATCTGGTCATCCTGG	Ehlers-Danlos syndrome, type 4
587779447	COL3A1	NM_000090.3(COL3A1):c.2842G>A (p.Gly948Arg)	CTTRGGATTGCTGGGATCACTGG	Ehlers-Danlos syndrome, type 4
587779456	COL3A1	NM_000090.3(COL3A1):c.2978G>A (p.Gly993Asp)	CGTGRTCCCCCTGGACCCCAGGG, ACGTGRTCCCCCTGGACCCCAGG	Ehlers-Danlos syndrome, type 4
587779466	COL3A1	NM_000090.3(COL3A1):c.2564G>A (p.Gly855Asp)	CCTGRTCCCCAAGGTGTCAAAGG	Ehlers-Danlos syndrome, type 4
587779472	COL3A1	NM_000090.3(COL3A1):c.3473G>A (p.Gly1158Asp)	CAGRTCCCATTGGACCACCAGGG, CCAGRTCCCATTGGACCACCAGG	Ehlers-Danlos syndrome, type 4
587779474	COL3A1	NM_000090.3(COL3A1):c.2068G>A (p.Gly690Arg)	GCARGGGCCCCAGGACTTAGAGG	Ehlers-Danlos syndrome, type 4
587779476	COL3A1	NM_000090.3(COL3A1):c.1466G>A (p.Gly489Glu)	CCTGRGTTCCGAGGACCTGCTGG	Ehlers-Danlos syndrome, type 4
587779478	COL3A1	NM_000090.3(COL3A1):c.809G>A (p.Gly270Glu)	GATGRACGAAATGGAGAAAAGGG, CGATGRACGAAATGGAGAAAAGG	Ehlers-Danlos syndrome, type 4

587779482	COL3A1	NM_000090.3(COL3A1):c.3508G>A (p.Gly1170Ser)	CAGARGTGAAAGAGGATCTGAGG	Ehlers-Danlos syndrome, type 4
587779484	COL3A1	NM_000090.3(COL3A1):c.2203G>A (p.Gly735Arg)	CTTRGAAGTCCTGGTCCAAAGGG, TCTTRGAAGTCCTGGTCCAAAGG	Ehlers-Danlos syndrome, type 4
587779493	COL3A1	NM_000090.3(COL3A1):c.1979G>A (p.Gly660Asp)	CAGGRTCCAAAGGGTGATGCCGG	Ehlers-Danlos syndrome, type 4
587779494	COL3A1	NM_000090.3(COL3A1):c.2555G>A (p.Gly852Asp)	TAGGRTCCTCCTGGTCCCCAAGG	Ehlers-Danlos syndrome, type 4
587779495	COL3A1	NM_000090.3(COL3A1):c.3437G>A (p.Gly1146Glu)	AGTGRACCTCCTGGCAAAGATGG	Ehlers-Danlos syndrome, type 4
587779499	COL3A1	NM_000090.3(COL3A1):c.1087G>A (p.Gly363Ser)	AATRGTGCCCCTGGACAAAGAGG	Ehlers-Danlos syndrome, type 4
587779501	COL3A1	NM_000090.3(COL3A1):c.3255+5G>A (p.Gly1068_Pro1085del)	CTGTAARTTTTGTCATTTTTTGG	Ehlers-Danlos syndrome, type 4
587779504	COL3A1	NM_000090.3(COL3A1):c.3562G>A (p.Gly1188Arg)	CCTRGACCTCCTGGTGCCCCTGG	Ehlers-Danlos syndrome, type 4
587779505	COL3A1	NM_000090.3(COL3A1):c.2708G>A (p.Gly903Glu)	GATGRGCCCCCAGGTCCTGCGGG, GGATGRGCCCCCAGGTCCTGCGG	Ehlers-Danlos syndrome, type 4
587779511	COL3A1	NM_000090.3(COL3A1):c.2888G>A (p.Gly963Asp)	CCAGRCATGCCAGGTCCTAGGGG, ACCAGRCATGCCAGGTCCTAGGG, CACCAGRCATGCCAGGTCCTAGG	Ehlers-Danlos syndrome, type 4
587779517	COL3A1	NM_000090.3(COL3A1):c.2825G>A (p.Gly942Glu)	AGGRAGCTCCAGGCCCACTTGGG, TAGGRAGCTCCAGGCCCACTTGG	Ehlers-Danlos syndrome, type 4
587779526	COL3A1	NM_000090.3(COL3A1):c.2510G>A (p.Gly837Asp)	GGAGRCCCTCCTGGAGTTGCAGG	Ehlers-Danlos syndrome, type 4
587779536	COL3A1	NM_000090.3(COL3A1):c.3391G>A (p.Gly1131Ser)	ATCRGCAGTCCAGGACCTGCAGG	Ehlers-Danlos syndrome, type 4
587779538	COL3A1	NM_000090.3(COL3A1):c.1149+5G>A (p.Gly351_Pro383del)	TGTAARTATCATAGTTGAGAGGG, CTGTAARTATCATAGTTGAGAGG	Ehlers-Danlos syndrome, type 4
587779540	COL3A1	NM_000090.3(COL3A1):c.3167G>A (p.Gly1056Asp)	GTCGRTCCAGCTGGAAAGAGTGG	Ehlers-Danlos syndrome, type 4
587779543	COL3A1	NM_000090.3(COL3A1):c.2185G>A (p.Gly729Arg)	CCTRGAGAAAGAGGAGGTCTTGG	Ehlers-Danlos syndrome, type 4

587779545	COL3A1	NM_000090.3(COL3A1):c.3140G>A (p.Gly1047Asp)	CCTGRTCATCCAGGCCCACCTGG	Ehlers-Danlos syndrome, type 4
587779549	COL3A1	NM_000090.3(COL3A1):c.2150G>A (p.Gly717Asp)	CCTGRTGCTGCTGGTACTCCTGG	Ehlers-Danlos syndrome, type 4
587779554	COL3A1	NM_000090.3(COL3A1):c.3220G>A (p.Gly1074Ser)	GCTRGTGCTCCCGGTCCTGCTGG	Ehlers-Danlos syndrome, type 4
587779557	COL3A1	NM_000090.3(COL3A1):c.637G>A (p.Gly213Ser)	CAGRGCCCTCCAGGACCTCCTGG	Ehlers-Danlos syndrome, type 4
587779561	COL3A1	NM_000090.3(COL3A1):c.3319G>A (p.Gly1107Arg)	AAARGACATCGAGGATTCCCTGG	Ehlers-Danlos syndrome, type 4
587779567	COL3A1	NM_000090.3(COL3A1):c.2833G>A (p.Gly945Ser)	CCARGCCCACTTGGGATTGCTGG	Ehlers-Danlos syndrome, type 4
587779569	COL3A1	NM_000090.3(COL3A1):c.1124G>A (p.Gly375Glu)	CAGGRACACGCTGGTGCTCAAGG	Ehlers-Danlos syndrome, type 4
587779576	COL3A1	NM_000090.3(COL3A1):c.2987G>A (p.Gly996Glu)	CCTGRACCCCAGGGTCTTCCTGG	Ehlers-Danlos syndrome, type 4
587779580	COL3A1	NM_000090.3(COL3A1):c.2905G>A (p.Gly969Arg)	AGGRGAAGCCCTGGCCCTCAGGG, TAGGRGAAGCCCTGGCCCTCAGG	Ehlers-Danlos syndrome, type 4
587779581	COL3A1	NM_000090.3(COL3A1):c.2168G>A (p.Gly723Asp)	CCTGRTCTGCAAGGAATGCCTGG	Ehlers-Danlos syndrome, type 4
587779583	COL3A1	NM_000090.3(COL3A1):c.2959G>A (p.Gly987Ser)	AACRGTCTCAGTGGAGAACGTGG	Thoracic aortic aneurysms and aortic dissections, Ehlers-Danlos syndrome, type 4
587779584	COL3A1	NM_000090.3(COL3A1):c.1618G>A (p.Gly540Arg)	CCCRGAAGTCCAGGAGGACCAGG	Thoracic aortic aneurysms and aortic dissections, Ehlers-Danlos syndrome, type 4
587779586	COL3A1	NM_000090.3(COL3A1):c.1268G>A (p.Gly423Asp)	AATGRTGCTCCTGGACTGCGAGG	Ehlers-Danlos syndrome, type 4
587779591	COL3A1	NM_000090.3(COL3A1):c.2087G>A (p.Gly696Asp)	AGAGRTGGAGCTGGTCCCCCTGG	Ehlers-Danlos syndrome, type 4
587779593	COL3A1	NM_000090.3(COL3A1):c.836G>A (p.Gly279Asp)	AACAGRTGCTCCTGGATTAAAGG	Ehlers-Danlos syndrome, type 4
587779595	COL3A1	NM_000090.3(COL3A1):c.2933G>A (p.Gly978Asp)	CAGGRTGAAAGTGGGAAACCAGG	Ehlers-Danlos syndrome, type 4
587779596	COL3A1	NM_000090.3(COL3A1):c.647G>A (p.Gly216Glu)	CCAGRACCTCCTGGTGCTATAGG	Ehlers-Danlos syndrome, type 4
587779599	COL3A1	NM_000090.3(COL3A1):c.2699G>A (p.Gly900Asp)	CCAGRCAAGGATGGGCCCCCAGG	Ehlers-Danlos syndrome, type 4
587779601	COL3A1	NM_000090.3(COL3A1):c.592G>A (p.Gly198Arg)	CCARGATACCAAGGACCCCCTGG	Ehlers-Danlos syndrome, type 4

587779606	COL3A1	NM_000090.3(COL3A1):c.2194G>A (p.Gly732Arg)	AGARGAGGTCTTGGAAGTCCTGG	Ehlers-Danlos syndrome, type 4
587779610	COL3A1	NM_000090.3(COL3A1):c.3284G>A (p.Gly1095Asp)	AAAGRTGAAACAGGTGAACGTGG	Ehlers-Danlos syndrome, type 4
587779611	COL3A1	NM_000090.3(COL3A1):c.1898G>A (p.Gly633Glu)	ACAGRACCCCCTGGTCCACAAGG	Ehlers-Danlos syndrome, type 4
587779621	COL3A1	NM_000090.3(COL3A1):c.1358G>A (p.Gly453Asp)	GCTGRTATTCCAGGTGTTCCAGG	Ehlers-Danlos syndrome, type 4
587779625	COL3A1	NM_000090.3(COL3A1):c.709G>A (p.Gly237Arg)	CCCRGACGACCTGGAGAGCGAGG	Thoracic aortic aneurysms and aortic dissections, Ehlers-Danlos syndrome, type 4
587779626	COL3A1	NM_000090.3(COL3A1):c.611G>A (p.Gly204Asp)	CCTGRTGAACCTGGGCAAGCTGG	Ehlers-Danlos syndrome, type 4
587779630	COL3A1	NM_000090.3(COL3A1):c.2293G>A (p.Gly765Ser)	ACTRGTCCTATTGGTCCTCCTGG	Ehlers-Danlos syndrome, type 4
587779631	COL3A1	NM_000090.3(COL3A1):c.1267G>A (p.Gly423Ser)	AATRGTGCTCCTGGACTGCGAGG	Ehlers-Danlos syndrome, type 4
587779633	COL3A1	NM_000090.3(COL3A1):c.1384G>A (p.Gly462Ser)	AAARGCGAAGATGGCAAGGATGG, AGCTAAARGCGAAGATGGCAAGG	Ehlers-Danlos syndrome, type 4
587779634	COL3A1	NM_000090.3(COL3A1):c.1844G>A (p.Gly615Glu)	CTGRACCTCAGGGACCCCCAGGG, ACTGRACCTCAGGGACCCCCAGG	Ehlers-Danlos syndrome, type 4
587779637	COL3A1	NM_000090.3(COL3A1):c.1249G>A (p.Gly417Arg)	CCARGACCAGCCGGTGCTAATGG	Ehlers-Danlos syndrome, type 4
587779638	COL3A1	NM_000090.3(COL3A1):c.2176G>A (p.Gly726Arg)	CAARGAATGCCTGGAGAAAGAGG	Ehlers-Danlos syndrome, type 4
587779641	COL3A1	NM_000090.3(COL3A1):c.593G>A (p.Gly198Glu)	CCAGRATACCAAGGACCCCCTGG	Ehlers-Danlos syndrome, type 4
587779642	COL3A1	NM_000090.3(COL3A1):c.2501G>A (p.Gly834Asp)	AAAGRTGAAGGAGGCCCTCCTGG	Ehlers-Danlos syndrome, type 4
587779644	COL3A1	NM_000090.3(COL3A1):c.827G>A (p.Gly276Asp)	AAGGRTGAAACAGGTGCTCCTGG	Ehlers-Danlos syndrome, type 4
587779648	COL3A1	NM_000090.3(COL3A1):c.3419G>A (p.Gly1140Glu)	TAGGRACCTGTTGGACCCAGTGG	Ehlers-Danlos syndrome, type 4
587779650	COL3A1	NM_000090.3(COL3A1):c.970G>A (p.Gly324Ser)	GACRGTGCTCGAGGCAGTGATGG	Ehlers-Danlos syndrome, type 4
587779656	COL3A1	NM_000090.3(COL3A1):c.701G>A (p.Gly234Asp)	TCAGRTAGACCCGGACGACCTGG	Ehlers-Danlos syndrome, type 4

587779662	COL3A1	NM_000090.3(COL3A1):c.2753G>A (p.Gly918Glu)	CCTGRAGTGTCTGGACCAAAAGG	Ehlers-Danlos syndrome, type 4
587779672	COL3A1	NM_000090.3(COL3A1):c.3266G>A (p.Gly1089Asp)	CAAGRCCCACGTGGTGACAAAGG	Ehlers-Danlos syndrome, type 4
587779673	COL3A1	NM_000090.3(COL3A1):c.998G>A (p.Gly333Asp)	CAGGRCCCTCCTGGTCCTCCTGG	Ehlers-Danlos syndrome, type 4
587779674	COL3A1	NM_000090.3(COL3A1):c.2860G>A (p.Gly954Arg)	ACTRGAGCACGGGGTCTTGCAGG	Ehlers-Danlos syndrome, type 4
587779678	COL3A1	NM_000090.3(COL3A1):c.2141G>A (p.Gly714Glu)	CCTGRGCCACCTGGTGCTGCTGG	Ehlers-Danlos syndrome, type 4
587779680	COL3A1	NM_000090.3(COL3A1):c.2186G>A (p.Gly729Glu)	CCTGRAGAAAGAGGAGGTCTTGG	Ehlers-Danlos syndrome, type 4
587779683	COL3A1	NM_000090.3(COL3A1):c.3544G>A (p.Gly1182Arg)	CCARGGCAACCAGGCCCTCCTGG	Ehlers-Danlos syndrome, type 4
587779689	COL3A1	NM_000090.3(COL3A1):c.2402G>A (p.Gly801Asp)	AGAGRTGAAACTGGCCCTCCAGG	Ehlers-Danlos syndrome, type 4
587779691	COL3A1	NM_000090.3(COL3A1):c.1763G>A (p.Gly588Asp)	TAGGRTGCTCCTGGTAAGAATGG	Ehlers-Danlos syndrome, type 4
587779692	COL3A1	NM_000090.3(COL3A1):c.1258G>A (p.Gly420Ser)	GCCRGTGCTAATGGTGCTCCTGG	Ehlers-Danlos syndrome, type 4
587779693	COL3A1	NM_000090.3(COL3A1):c.1556G>A (p.Gly519Glu)	AGAGRAGCTGCTGGAGAACCTGG	Ehlers-Danlos syndrome, type 4
587779695	COL3A1	NM_000090.3(COL3A1):c.2131G>A (p.Gly711Ser)	GCTRGTCCTCCTGGGCCACCTGG	Ehlers-Danlos syndrome, type 4
587779696	COL3A1	NM_000090.3(COL3A1):c.1096G>A (p.Gly366Arg)	CCTRGACAAAGAGGAGAACCTGG	Ehlers-Danlos syndrome, type 4
587779698	COL3A1	NM_000090.3(COL3A1):c.2177G>A (p.Gly726Glu)	CAAGRAATGCCTGGAGAAAGAGG	Ehlers-Danlos syndrome, type 4
587779706	COL3A1	NM_000090.3(COL3A1):c.2096G>A (p.Gly699Asp)	GCTGRTCCCCCTGGTCCCGAAGG	Ehlers-Danlos syndrome, type 4
587779711	COL3A1	NM_000090.3(COL3A1):c.610G>A (p.Gly204Ser)	CCTRGTGAACCTGGGCAAGCTGG	Ehlers-Danlos syndrome, type 4
587779714	COL3A1	NM_000090.3(COL3A1):c.539G>A (p.Gly180Asp)	CCAGRCCCTCCCGGTCCCCCTGG	Ehlers-Danlos syndrome, type 4
587779716	COL3A1	NM_000090.3(COL3A1):c.2735G>A (p.Gly912Asp)	ACTGRTGCTCCTGGCAGCCCTGG	Ehlers-Danlos syndrome, type 4
587779718	COL3A1	NM_000090.3(COL3A1):c.799G>A (p.Gly267Ser)	TAGRGCTTCGATGGACGAAATGG	Ehlers-Danlos syndrome, type 4
587779723	COL3A1	NM_000090.3(COL3A1):c.	CCCTRGCCCTCAGGGTGTCAAGG	Ehlers-Danlos syndrome, type 4

		2914G>A (p.Gly972Ser)		
112456072	COL3A1	NM_000090.3(COL3A1):c.3563G>A (p.Gly1188Glu)	CCTGRACCTCCTGGTGCCCCTGG	Ehlers-Danlos syndrome, type 4
794728060	COL3A1	NM_000090.3(COL3A1):c.4087C>T (p.Arg1363Ter)	CCTTCGACTTCTCTCCAGCYGAG	Thoracic aortic aneurysms and aortic dissections
587779527	COL3A1	NM_000090.3(COL3A1):c.1786C>T (p.Arg596Ter)	CCTGGTAAGAATGGAGAAYGAGG	Ehlers-Danlos syndrome, type 4
672601346	COL4A1	NM_001845.5(COL4A1):c.2263G>A (p.Gly755Arg)	CCCRGGGAGAAGGGGAGCATTGG	Brain small vessel disease with hemorrhage
672601349	COL4A1	NM_001845.5(COL4A1):c.2122G>A (p.Gly708Arg)	ATGRGGCCACCGGGGACTCCAGG	Brain small vessel disease with hemorrhage
121912857	COL4A1	NM_001845.5(COL4A1):c.1685G>A (p.Gly562Glu)	CCTGRAAGAGATGGCCATCCGGG, TCCTGRAAGAGATGGCCATCCGG	Brain small vessel disease with hemorrhage
606231465	COL4A1	NM_001845.5(COL4A1):c.2194-1G>A	TTTCARGGAGAGCCTGGAGTTGG	Brain small vessel disease with hemorrhage
113994105	COL4A1	NM_001845.5(COL4A1):c.1555G>A (p.Gly519Arg)	CCARGGCTGATAGGCCAGCCAGG	Angiopathy, hereditary, with nephropathy, aneurysms, and muscle cramps
113994107	COL4A1	NM_001845.5(COL4A1):c.1769G>A (p.Gly590Glu)	CCTGGAGRAGTTGGATTCCCAGG	Brain small vessel disease with hemorrhage
113994108	COL4A1	NM_001845.5(COL4A1):c.2159G>A (p.Gly720Asp)	AATGRCTTACCTGGGAACCCAGG	Brain small vessel disease with hemorrhage
113994109	COL4A1	NM_001845.5(COL4A1):c.2245G>A (p.Gly749Ser)	CCCRGCATTCCTGGCACACCCGG	Familial porencephaly
113994111	COL4A1	NM_001845.5(COL4A1):c.3389G>A (p.Gly1130Asp)	CCTGRTGTCAAAGGAGAAGCAGG	Familial porencephaly
113994112	COL4A1	NM_001845.5(COL4A1):c.3706G>A (p.Gly1236Arg)	AAARGAGACCGCGGACCTCAGGG, CAAARGAGACCGCGGACCTCAGG	Familial porencephaly
587777379	COL4A1	NM_001845.5(COL4A1):c.3976G>A (p.Gly1326Arg)	CCTTGATCACCTTTAATTCYCTG	Schizencephaly
387906602	COL4A2	NM_001846.2(COL4A2):c.3455G>A (p.Gly1152Asp)	CAGRCTTTCCAGGGCTGACTGGG, CCAGRCTTTCCAGGGCTGACTGG	Porencephaly 2
387906603	COL4A2	NM_001846.2(COL4A2):c.3110G>A (p.Gly1037Glu)	AAGGRAGACATCGGAGTCCCCGG	Porencephaly 2
121912858	COL4A4	NM_000092.4(COL4A4):c.3601G>A (p.Gly1201Ser)	CCTRGTCCAGTGGGAATACCTGG	Alport syndrome, autosomal recessive
121912860	COL4A4	NM_000092.4(COL4A4):c.	GATGRGCTACCTGGTCCTCCAGG	Benign familial hematuria

		2690G>A (p.Gly897Glu)		
281874656	COL4A5	NM_000495.4(COL4A5):c.1084G>A (p.Gly362Arg)	ATTRGGTTGCCTGGGTTGCCTGG	Alport syndrome, X-linked recessive
281874660	COL4A5	NM_000495.4(COL4A5):c.1216G>A (p.Gly406Ser)	AGGRGTCAGAAAGGTGATGAAGG	Alport syndrome, X-linked recessive
281874663	COL4A5	NM_000495.4(COL4A5):c.1259G>A (p.Gly420Glu)	CCTGRACCTCCTGGACTTGACGG	Alport syndrome, X-linked recessive
281874664	COL4A5	NM_000495.4(COL4A5):c.1294G>A (p.Gly432Arg)	CCTRGGCTTCCAGGGCCTCCTGG	Alport syndrome, X-linked recessive
281874669	COL4A5	NM_000495.4(COL4A5):c.142G>A (p.Gly48Arg)	CAGRGAGAGAGAGGGTTTCCAGG	Alport syndrome, X-linked recessive
281874671	COL4A5	NM_000495.4(COL4A5):c.1589G>A (p.Gly530Asp)	CAGGRCATTCCAGGAGCTCCAGG	Alport syndrome, X-linked recessive
281874672	COL4A5	NM_000495.4(COL4A5):c.1598G>A (p.Gly533Glu)	CCAGRAGCTCCAGGTGCTCCAGG	Alport syndrome, X-linked recessive
281874675	COL4A5	NM_000495.4(COL4A5):c.1726G>A (p.Gly576Ser)	CCTRGCACTCCTGGACAGGATGG, TTTACCTRGCACTCCTGGACAGG	Alport syndrome, X-linked recessive
281874677	COL4A5	NM_000495.4(COL4A5):c.1744G>A (p.Gly582Arg)	GATRGATTGCCAGGGCTTCCTGG	Alport syndrome, X-linked recessive
281874680	COL4A5	NM_000495.4(COL4A5):c.1835G>A (p.Gly612Asp)	CAGRCCTCCCAGGGAATATAGGG, CCAGRCCTCCCAGGGAATATAGG	Alport syndrome, X-linked recessive
281874683	COL4A5	NM_000495.4(COL4A5):c.1904G>A (p.Gly635Asp)	AAAGRCATACAAGGTGTGGCAGG, TGAAAAAGRCATACAAGGTGTGG	Alport syndrome, X-linked recessive
281874689	COL4A5	NM_000495.4(COL4A5):c.2288G>A (p.Gly763Glu)	CCAGRACTTCCAGGTTTCAAAGG	Alport syndrome, X-linked recessive
281874690	COL4A5	NM_000495.4(COL4A5):c.2305G>A (p.Gly769Arg)	AAARGAGCACTTGGTCCAAAAGG	Alport syndrome, X-linked recessive
281874695	COL4A5	NM_000495.4(COL4A5):c.2483G>A (p.Gly828Glu)	CCAGRGATTCCTGGGCCAATAGG	Alport syndrome, X-linked recessive
281874703	COL4A5	NM_000495.4(COL4A5):c.2722G>A (p.Gly908Arg)	CCARGACCTTTGGGAATTCCTGG	Alport syndrome, X-linked recessive
281874704	COL4A5	NM_000495.4(COL4A5):c.2731G>A (p.Gly911Arg)	TTGRGAATTCCTGGCAGGAGTGG, GACCTTTGRGAATTCCTGGCAGG	Alport syndrome, X-linked recessive
281874706	COL4A5	NM_000495.4(COL4A5):c.286G>A (p.Gly96Arg)	CCTRGACTTCCTGGATTTCCAGG	Alport syndrome, X-linked recessive
281874717	COL4A5	NM_000495.4(COL4A5):c.3587G>A (p.Gly1196Glu)	CCTGRACTTCCAGGACTTTCTGG	Alport syndrome, X-linked recessive

281874722	COL4A5	NM_000495.4(COL4A5):c.385G>A (p.Gly129Arg)	TAGRGAGAACGTGGATTTCCAGG	Alport syndrome, X-linked recessive
281874725	COL4A5	NM_000495.4(COL4A5):c.3925-1G>A	TCARGGTAATCCTGGCCGGCCGG, TTATTCARGGTAATCCTGGCCGG	Alport syndrome, X-linked recessive
281874733	COL4A5	NM_000495.4(COL4A5):c.4271G>A (p.Gly1424Glu)	AAAGRAGACCCAGGTCTGCCAGG	Alport syndrome, X-linked recessive
281874739	COL4A5	NM_000495.4(COL4A5):c.438+5G>A	TCCAGTAARTTATAAAATTTGGG	Alport syndrome, X-linked recessive
281874746	COL4A5	NM_000495.4(COL4A5):c.4702G>A (p.Glu1568Lys)	AGTATGTRAAGCTCCAGCTGTGG	Alport syndrome, X-linked recessive
281874763	COL4A5	NM_000495.4(COL4A5):c.689G>A (p.Gly230Asp)	CAGGRTGAGCAAGGTCTTCAGGG , ACAGGRTGAGCAAGGTCTTCAGG	Alport syndrome, X-linked recessive
104886080	COL4A5	NM_000495.4(COL4A5):c.892G>A (p.Gly298Ser)	TAGRGTAAACCAGGCAAAGATGG	Alport syndrome, X-linked recessive
104886381	COL4A5	NM_000495.4(COL4A5):c.3554-1G>A	CTGACARGTCAACCAGGCTTTGG	Alport syndrome, X-linked recessive
587776402	COL4A5	NM_000495.4(COL4A5):c.4199-1G>A	GTARGTCCAACTGGCCCTCCAGG	Alport syndrome, X-linked recessive
104886043	COL4A5	NM_000495.4(COL4A5):c.161G>A (p.Gly54Asp)	CCAGRTTTGGAAGGACACCCAGG	Alport syndrome, X-linked recessive
104886057	COL4A5	NM_000495.4(COL4A5):c.593G>A (p.Gly198Glu)	CCCAGRACCACCAGGTTTGATGG	Alport syndrome, X-linked recessive
104886060	COL4A5	NM_000495.4(COL4A5):c.574G>A (p.Gly192Arg)	CCARGGCCAATTGGTCCCCCAGG	Alport syndrome, X-linked recessive
104886061	COL4A5	NM_000495.4(COL4A5):c.584G>A (p.Gly195Asp)	ATTGRTCCCCCAGGACCACCAGG	Alport syndrome, X-linked recessive
104886070	COL4A5	NM_000495.4(COL4A5):c.791G>A (p.Gly264Asp)	CCTGRTGACCGAGGGCCTCCTGG	Alport syndrome, X-linked recessive
104886075	COL4A5	NM_000495.4(COL4A5):c.655G>A (p.Gly219Ser)	AATATGRGCTTAAATTTCCAGGG, GAATATGRGCTTAAATTTCCAGG	Alport syndrome, X-linked recessive
104886084	COL4A5	NM_000495.4(COL4A5):c.937G>A (p.Gly313Ser)	AAGRGTTTGCCTGGTGATCCTGG	Alport syndrome, X-linked recessive
104886086	COL4A5	NM_000495.4(COL4A5):c.956G>A (p.Gly319Asp)	CCTGRTTACCCTGGTGAACCCGG	Alport syndrome, X-linked recessive
104886091	COL4A5	NM_000495.4(COL4A5):c.974G>A (p.Gly325Glu)	ACCCGRAAGGGATGGTGAAAAGG	Alport syndrome, X-linked recessive
104886096	COL4A5	NM_000495.4(COL4A5):c.1094G>A (p.Gly365Glu)	CCTGRGTTGCCTGGAGAAAAAGG	Alport syndrome, X-linked recessive

104886097	COL4A5	NM_000495.4(COL4A5):c.1112G>A (p.Gly371Glu)	AAAGRAGAGCGAGGATTTCCTGG	Alport syndrome, X-linked recessive
104886098	COL4A5	NM_000495.4(COL4A5):c.1139G>A (p.Gly380Asp)	CAGGRTCCACCTGGCCTTCCTGG	Alport syndrome, X-linked recessive
104886101	COL4A5	NM_000495.4(COL4A5):c.1226G>A (p.Gly409Asp)	AAAGRTGATGAAGGACCACCTGG	Alport syndrome, X-linked recessive
104886103	COL4A5	NM_000495.4(COL4A5):c.1243G>A (p.Gly415Arg)	CCACCTRGAATTTCCATTCCTGG	Alport syndrome, X-linked recessive
104886105	COL4A5	NM_000495.4(COL4A5):c.1148G>A (p.Gly383Asp)	CCTGRCCTTCCTGGACCTCCAGG	Alport syndrome, X-linked recessive
104886107	COL4A5	NM_000495.4(COL4A5):c.1199G>A (p.Gly400Glu)	CCTGRATTTCCTGGAGAAAGGGG, TCCTGRATTTCCTGGAGAAAGGG, CTCCTGRATTTCCTGGAGAAAGG	Alport syndrome, X-linked recessive
104886110	COL4A5	NM_000495.4(COL4A5):c.1268G>A (p.Gly423Glu)	CTGRACTTGACGGACAGCCTGGG, CCTGRACTTGACGGACAGCCTGG	Alport syndrome, X-linked recessive
104886111	COL4A5	NM_000495.4(COL4A5):c.1276G>A (p.Gly426Arg)	GACRGACAGCCTGGGGCTCCTGG	Alport syndrome, X-linked recessive
104886112	COL4A5	NM_000495.4(COL4A5):c.1286G>A (p.Gly429Glu)	CTGRGGCTCCTGGGCTTCCAGGG, CCTGRGGCTCCTGGGCTTCCAGG	Alport syndrome, X-linked recessive
104886114	COL4A5	NM_000495.4(COL4A5):c.1397G>A (p.Gly466Glu)	AAAGRACTCCAAGGAGAACAAGG	Alport syndrome, X-linked recessive
104886115	COL4A5	NM_000495.4(COL4A5):c.1406G>A (p.Gly469Glu)	CAAGRAGAACAAGGAGTGAAAGG	Alport syndrome, X-linked recessive
104886117	COL4A5	NM_000495.4(COL4A5):c.1472G>A (p.Gly491Glu)	TCAGRGCCTCCAGGTCAACCTGG	Alport syndrome, X-linked recessive
104886118	COL4A5	NM_000495.4(COL4A5):c.1481G>A (p.Gly494Asp)	CCAGRTCAACCTGGTTTGCCAGG	Alport syndrome, X-linked recessive
104886122	COL4A5	NM_000495.4(COL4A5):c.1562G>A (p.Gly521Asp)	GCTGRTGCAACTGGTCCCAAAGG	Alport syndrome, X-linked recessive
104886125	COL4A5	NM_000495.4(COL4A5):c.1607G>A (p.Gly536Asp)	CCAGRTGCTCCAGGCTTTCCTGG	Alport syndrome, X-linked recessive
104886130	COL4A5	NM_000495.4(COL4A5):c.1736G>A (p.Gly579Glu)	CTGRACAGGATGGATTGCCAGGG, CCTGRACAGGATGGATTGCCAGG	Alport syndrome, X-linked recessive
104886131	COL4A5	NM_000495.4(COL4A5):c.1780G>A (p.Gly594Ser)	TAGRGTGGAATTACTTTTAAGGG, TTAGRGTGGAATTACTTTTAAGG	Alport syndrome, X-linked recessive
104886132	COL4A5	NM_000495.4(COL4A5):c.	TAGGGTRGAATTACTTTTAAGGG,	Alport syndrome, X-linked recessive

		1783G>A (p.Gly595Arg)	TTAGGGTRGAATTACTTTTAAGG	
104886136	COL4A5	NM_000495.4(COL4A5):c.1681G>A (p.Gly561Arg)	AAARGAGAGTTGGGTTCCCCTGG	Alport syndrome, X-linked recessive
104886138	COL4A5	NM_000495.4(COL4A5):c.1718G>A (p.Gly573Asp)	CCTGRTTTACCTGGCACTCCTGG	Alport syndrome, X-linked recessive
104886139	COL4A5	NM_000495.4(COL4A5):c.1735G>A (p.Gly579Arg)	CCTRGACAGGATGGATTGCCAGG	Alport syndrome, X-linked recessive
104886142	COL4A5	NM_000495.4(COL4A5):c.1871G>A (p.Gly624Asp)	CCCCCTGRTTTCGGCCCTCCAGG	Alport syndrome, X-linked recessive
104886144	COL4A5	NM_000495.4(COL4A5):c.1886G>A (p.Gly629Asp)	CCAGRCCCAGTAGGTGAAAAAG G	Alport syndrome, X-linked recessive
104886145	COL4A5	NM_000495.4(COL4A5):c.1895G>A (p.Gly632Asp)	GTAGRTGAAAAAGGCATACAAG G	Alport syndrome, X-linked recessive
104886146	COL4A5	NM_000495.4(COL4A5):c.1897G>A (p.Glu633Lys)	GTAGGTRAAAAAGGCATACAAG G	Alport syndrome, X-linked recessive
104886147	COL4A5	NM_000495.4(COL4A5):c.1912G>A (p.Gly638Ser)	CAARGTGTGGCAGGAAATCCAGG	Alport syndrome, X-linked recessive
104886157	COL4A5	NM_000495.4(COL4A5):c.2023G>A (p.Gly675Ser)	GATRGTGATGTAGGTCTTCCAGG	Alport syndrome, X-linked recessive
104886158	COL4A5	NM_000495.4(COL4A5):c.2042G>A (p.Gly681Asp)	TAGRTGACCCTGGACTTCCAGGG, ATAGRTGACCCTGGACTTCCAGG	Alport syndrome, X-linked recessive
104886163	COL4A5	NM_000495.4(COL4A5):c.2165G>A (p.Gly722Glu)	CAGRACCTCCAGGAGCACCTGGG , CCAGRACCTCCAGGAGCACCTGG	Alport syndrome, X-linked recessive
104886165	COL4A5	NM_000495.4(COL4A5):c.2219G>A (p.Gly740Glu)	CCTGRGCCACCCGGCTTTCCAGG	Alport syndrome, X-linked recessive
104886166	COL4A5	NM_000495.4(COL4A5):c.2228G>A (p.Gly743Asp)	ACCCGRCTTTCCAGGACCAAAGG	Alport syndrome, X-linked recessive
104886168	COL4A5	NM_000495.4(COL4A5):c.2060G>A (p.Gly687Glu)	CAGRGCAACCAGGCTTGCCAGGG , CCAGRGCAACCAGGCTTGCCAGG	Alport syndrome, X-linked recessive
104886171	COL4A5	NM_000495.4(COL4A5):c.2287G>A (p.Gly763Arg)	CCARGACTTCCAGGTTTCAAAGG	Alport syndrome, X-linked recessive
104886172	COL4A5	NM_000495.4(COL4A5):c.2297G>A (p.Gly766Asp)	CCAGRTTTCAAAGGAGCACTTGG	Alport syndrome, X-linked recessive
104886174	COL4A5	NM_000495.4(COL4A5):c.2332G>A (p.Gly778Ser)	CGTRGTTTCCCAGGACCTCCGGG, TCGTRGTTTCCCAGGACCTCCGG	Alport syndrome, X-linked recessive
104886177	COL4A5	NM_000495.4(COL4A5):c.2386G>A (p.Gly796Arg)	CCCTRGACCAAAAGGTATGGAGG ,	Alport syndrome, X-linked recessive

			GCTCCCTRGACCAAAAGGTATGG	
104886179	COL4A5	NM_000495.4(COL4A5):c.2404G>A (p.Gly802Arg)	GTTRGACCAAATGGACAACCTGG	Alport syndrome, X-linked recessive
104886180	COL4A5	NM_000495.4(COL4A5):c.2423G>A (p.Gly808Glu)	CTGRACCAATGGGACCTCCTGGG, CCTGRACCAATGGGACCTCCTGG	Alport syndrome, X-linked recessive
104886182	COL4A5	NM_000495.4(COL4A5):c.2431G>A (p.Gly811Arg)	ATGRGACCTCCTGGGCTGCCAGG	Alport syndrome, X-linked recessive
104886186	COL4A5	NM_000495.4(COL4A5):c.2554G>A (p.Gly852Arg)	CCTCCTRGACTTGATGTTCCAGG	Alport syndrome, X-linked recessive
104886187	COL4A5	NM_000495.4(COL4A5):c.2555G>A (p.Gly852Glu)	CCTCCTGRACTTGATGTTCCAGG	Alport syndrome, X-linked recessive
104886188	COL4A5	NM_000495.4(COL4A5):c.2597G>A (p.Gly866Glu)	CCAGRGATCCCCGGAGCACCTGG	Alport syndrome, X-linked recessive
104886189	COL4A5	NM_000495.4(COL4A5):c.2605G>A (p.Gly869Arg)	CCCRGAGCACCTGGTCCTATAGG	Alport syndrome, X-linked recessive
104886191	COL4A5	NM_000495.4(COL4A5):c.2624G>A (p.Gly875Glu)	TAGRACCTCCAGGATCACCAGGG , ATAGRACCTCCAGGATCACCAGG	Alport syndrome, X-linked recessive
104886195	COL4A5	NM_000495.4(COL4A5):c.2804G>A (p.Gly935Asp)	CCTGRCCCTACAGGAGAAAAAGG	Alport syndrome, X-linked recessive
104886210	COL4A5	NM_000495.4(COL4A5):c.3088G>A (p.Gly1030Ser)	ATCRGTGATATGGGTTTTCCAGG	Alport syndrome, X-linked recessive
104886214	COL4A5	NM_000495.4(COL4A5):c.3115G>A (p.Gly1039Ser)	CAGRGTGTGGAAGGGCCTCCTGG	Alport syndrome, X-linked recessive
104886215	COL4A5	NM_000495.4(COL4A5):c.3134G>A (p.Gly1045Glu)	CCTGRACCTTCTGGAGTTCCTGG	Alport syndrome, X-linked recessive
104886223	COL4A5	NM_000495.4(COL4A5):c.3247G>A (p.Gly1083Ser)	CAGRGTGAGCCTGGTCTGCCTGG	Alport syndrome, X-linked recessive
104886225	COL4A5	NM_000495.4(COL4A5):c.3319G>A (p.Gly1107Arg)	CCCRGATTACCAGGAACCCCTGG	Alport syndrome, X-linked recessive
104886228	COL4A5	NM_000495.4(COL4A5):c.3427G>A (p.Gly1143Ser)	CCCRGCCTTCCAGGAGAACCTGG	Alport syndrome, X-linked recessive
104886229	COL4A5	NM_000495.4(COL4A5):c.3428G>A (p.Gly1143Asp)	CCCGRCCTTCCAGGAGAACCTGG	Alport syndrome, X-linked recessive
104886232	COL4A5	NM_000495.4(COL4A5):c.3257G>A (p.Gly1086Asp)	CTGRTCTGCCTGGATACCCAGGG, CCTGRTCTGCCTGGATACCCAGG	Alport syndrome, X-linked recessive
104886235	COL4A5	NM_000495.4(COL4A5):c.3481G>A (p.Gly1161Arg)	CCARGGCCTCCAGGCGAAAAAG G	Alport syndrome, X-linked recessive
104886236	COL4A5	NM_000495.4(COL4A5):c.	AAARGCAAACCCGGTCAAGATG	Alport syndrome, X-linked recessive

		3499G>A (p.Gly1167Ser)	G	
104886237	COL4A5	NM_000495.4(COL4A5):c.3508G>A (p.Gly1170Ser)	CCCRGTCAAGATGGTATTCCTGG	Alport syndrome, X-linked recessive
104886240	COL4A5	NM_000495.4(COL4A5):c.3535G>A (p.Gly1179Arg)	GCTRGACAGAAGGGTGAACCAGG	Alport syndrome, X-linked recessive
104886244	COL4A5	NM_000495.4(COL4A5):c.3586G>A (p.Gly1196Arg)	CCTRGACTTCCAGGACTTTCTGG	Alport syndrome, X-linked recessive
104886247	COL4A5	NM_000495.4(COL4A5):c.3632G>A (p.Gly1211Glu)	CCTGRGATTCCAGGAAATCCTGG	Alport syndrome, X-linked recessive
104886248	COL4A5	NM_000495.4(COL4A5):c.3641G>A (p.Gly1214Glu)	CCAGRAAATCCTGGCCTTCCAGG	Alport syndrome, X-linked recessive
104886250	COL4A5	NM_000495.4(COL4A5):c.3694G>A (p.Gly1232Ser)	CCTRGTGTGCAGGGTCCCCCAGG	Alport syndrome, X-linked recessive
104886251	COL4A5	NM_000495.4(COL4A5):c.3659G>A (p.Gly1220Asp)	CCAGRTCCAAAGGGCGAACCAGG	Alport syndrome, X-linked recessive
104886253	COL4A5	NM_000495.4(COL4A5):c.3686G>A (p.Gly1229Asp)	CACGRTTTCCCTGGTGTGCAGGG, TCACGRTTTCCCTGGTGTGCAGG	Alport syndrome, X-linked recessive
104886257	COL4A5	NM_000495.4(COL4A5):c.3808G>A (p.Gly1270Ser)	GAARGTCCTCCAGGTCTCCCTGG	Alport syndrome, X-linked recessive
104886261	COL4A5	NM_000495.4(COL4A5):c.3731G>A (p.Gly1244Asp)	TCTCCGGRTCCAGCTCTGGAAGG	Alport syndrome, X-linked recessive
104886262	COL4A5	NM_000495.4(COL4A5):c.3754G>A (p.Gly1252Ser)	AAARGCAACCCTGGGCCCCAAGG	Alport syndrome, X-linked recessive
104886263	COL4A5	NM_000495.4(COL4A5):c.3763G>A (p.Gly1255Arg)	CCTRGGCCCCAAGGTCCTCCTGG	Alport syndrome, X-linked recessive
104886279	COL4A5	NM_000495.4(COL4A5):c.4342G>A (p.Gly1448Ser)	CAARGTCCCCCAGGTCCCCCTGG	Alport syndrome, X-linked recessive
104886297	COL4A5	NM_000495.4(COL4A5):c.4787G>A (p.Gly1596Asp)	GATTGRTTATTCCTTCATGATGG	Alport syndrome, X-linked recessive
104886331	COL4A5	NM_000495.4(COL4A5):c.1516+1G>A	TCCAGRTAAATTATGCCTCAGGG, CTCCAGRTAAATTATGCCTCAGG	Alport syndrome, X-linked recessive
104886338	COL4A5	NM_000495.4(COL4A5):c.1780-1G>A	TTARGGTGGAATTACTTTTAAGG	Alport syndrome, X-linked recessive
104886361	COL4A5	NM_000495.4(COL4A5):c.2705G>A (p.Gly902Glu)	ATGGRACCTCCAGGCCCACCAGG	Alport syndrome, X-linked recessive
104886363	COL4A5	NM_000495.4(COL4A5):c.2732G>A (p.Gly911Glu)	TTGGRAATTCCTGGCAGGAGTGG	Alport syndrome, X-linked recessive
104886370	COL4A5	NM_000495.4(COL4A5):c.2840G>A (p.Gly947Asp)	CCTGRCCTTCCAGGCCCTCCTGG	Alport syndrome, X-linked recessive

104886378	COL4A5	NM_000495.4(COL4A5):c.3017-1G>A	CTARGTCCCAAAGGTAACCCTGG	Alport syndrome, X-linked recessive
104886384	COL4A5	NM_000495.4(COL4A5):c.3605-1G>A	ATARGCCAAAAGGGTGATGGAGG, TTCATARGCCAAAAGGGTGATGG	Alport syndrome, X-linked recessive
104886396	COL4A5	NM_000495.4(COL4A5):c.385-719G>A	CAAGRTGGAGAGAAGGGTATTGG	Alport syndrome, X-linked recessive
794727397	COL4A5	NM_000495.4(COL4A5):c.1844G>A (p.Gly615Glu)	CCAGRGAATATAGGGCCTATGGG, CCCAGRGAATATAGGGCCTATGG	Alport syndrome, X-linked recessive
281874676	COL4A5	NM_000495.4(COL4A5):c.1738C>T (p.Gln580Ter)	CCTGGCACTCCTGGAYAGGATGG	Alport syndrome, X-linked recessive
281874681	COL4A5	NM_000495.4(COL4A5):c.1856C>T (p.Pro619Leu)	CCCAGGGAATATAGGGCYTATGG, CCAGGGAATATAGGGCYTATGGG	Alport syndrome, X-linked recessive
281874727	COL4A5	NM_000495.4(COL4A5):c.4147C>T (p.Gln1383Ter)	CCTCCAGGAATCCCTGGCYAGCC, CCAGGAATCCCTGGCYAGCCTGG	Alport syndrome, X-linked recessive
281874661	COL4A5	NM_000495.4(COL4A5):c.1219C>T (p.Gln407Ter)	CCTGGAGAAAGGGGTYAGAAAGG	Alport syndrome, X-linked recessive
104886094	COL4A5	NM_000495.4(COL4A5):c.1117C>T (p.Arg373Ter)	CCTGGAGAAAAAGGAGAGYGAGG	Alport syndrome, X-linked recessive
104886207	COL4A5	NM_000495.4(COL4A5):c.3046C>T (p.Gln1016Ter)	CCCTGGTCTCCCTGGAYAGCCAG, CCTGGTCTCCCTGGAYAGCCAGG	Alport syndrome, X-linked recessive
104886213	COL4A5	NM_000495.4(COL4A5):c.3181C>T (p.Gln1061Ter)	CCCCAGGATTACCTGGAYAGAAA, CCCAGGATTACCTGGAYAGAAAG, CCAGGATTACCTGGAYAGAAAGG	Alport syndrome, X-linked recessive
104886241	COL4A5	NM_000495.4(COL4A5):c.3538C>T (p.Gln1180Ter)	CCTGGACCAGCTGGAYAGAAGGG	Alport syndrome, X-linked recessive
104886270	COL4A5	NM_000495.4(COL4A5):c.4228C>T (p.Arg1410Cys)	CCCTCCAGGAGATCCTGGAYGCA, CCTCCAGGAGATCCTGGAYGCAA, CCAGGAGATCCTGGAYGCAATGG	Alport syndrome, X-linked recessive
104886286	COL4A5	NM_000495.4(COL4A5):c.4687C>T (p.Arg1563Ter)	CCAGCCATTCATTAGTYGGTAAG	Alport syndrome, X-linked recessive
61735045	COL5A1	NM_000093.4(COL5A1):c.1588G>A (p.Gly530Ser)	GGCRGCGATGCGGGCTCCAAAGG	Ehlers-Danlos syndrome, classic type, not specified

121912935	COL6A1	NM_001848.2(COL6A1):c.1022G>A (p.Gly341Asp)	CCAGRCCTGCCAGGCTGCAAGGG, CCCAGRCCTGCCAGGCTGCAAGG	Bethlem myopathy
794727788	COL6A2	NM_001849.3(COL6A2):c.812G>A (p.Gly271Asp)	AAGGRCAACATGGGTGAGCCGGG, CAAGGRCAACATGGGTGAGCCGG	Ullrich congenital muscular dystrophy, Bethlem myopathy
121912940	COL6A2	NM_001849.3(COL6A2):c.811G>A (p.Gly271Ser)	AAGRGCAACATGGGTGAGCCGGG, CAAGRGCAACATGGGTGAGCCGG	Bethlem myopathy
727502827	COL6A2	NM_001849.3(COL6A2):c.857G>A (p.Gly286Glu)	CAGGRAGACCCGGGCATCGAAGG	Congenital muscular dystrophy
727502828	COL6A2	NM_001849.3(COL6A2):c.874G>A (p.Gly292Ser)	GAARGCCCCATTGGATTCCCAGG	Congenital muscular dystrophy
267606750	COL6A2	NM_001849.3(COL6A2):c.1861G>A (p.Asp621Asn)	ATCRACAGCTCCGAGAGCATTGG	Congenital muscular dystrophy, Bethlem myopathy
398123646	COL6A2	NM_001849.3(COL6A2):c.1522-1G>A	TCCTCCARGGAGACCCCGGCAGG	Ullrich congenital muscular dystrophy, Bethlem myopathy, not provided
376880198	COL6A2	NM_001849.3(COL6A2):c.2527C>T (p.Arg843Trp)	CCTGCTGGACGGCTCCGAGYGGC	
121912942	COL6A2	NM_001849.3(COL6A2):c.2455C>T (p.Gln819Ter)	CCCAGACCTTCCCTGCYAAACAG, CCAGACCTTCCCTGCYAAACAGG	Myosclerosis, autosomal recessive
117725825	COL6A2	NM_001849.3(COL6A2):c.2795C>T (p.Pro932Leu)	CCATCGTGCGCAGCCYGCGTGGC	Congenital muscular dystrophy, Bethlem myopathy, not provided
794727188	COL6A3	NM_004369.3(COL6A3):c.6239G>A (p.Gly2080Asp)	AACGRCACTCAAGGTTTCCAGGG, GAACGRCACTCAAGGTTTCCAGG	Bethlem myopathy
398124128	COL6A3	NM_004369.3(COL6A3):c.6282+1G>A	AAGRTGAGGCGTGGGTGATGGGG, AAAGRTGAGGCGTGGGTGATGGG, TAAAGRTGAGGCGTGGGTGATGG	not provided
121434554	COL6A3	NM_004369.3(COL6A3):c.1393C>T (p.Arg465Ter)	CCAACTTCAATGCCATCYGAGAC	Ullrich congenital muscular dystrophy
121912829	COL7A1	NM_000094.3(COL7A1):c.6118G>A (p.Gly2040Ser)	CCTRGTATTCCCGGGCTCCCAGG	Generalized dominant dystrophic epidermolysis bullosa
121912832	COL7A1	NM_000094.3(COL7A1):c.6007G>A (p.Gly2003Arg)	CGCRGGCTGAAGGGCGACCGTGG	Dominant dystrophic epidermolysis bullosa with absence of skin

121912836	COL7A1	NM_000094.3(COL7A1):c.6127G>A (p.Gly2043Arg)	CCCRGGCTCCCAGGCAGGGCTGG, TATTCCCRGGCTCCCAGGCAGGG, GTATTCCCRGGCTCCCAGGCAGG	Generalized dominant dystrophic epidermolysis bullosa
121912837	COL7A1	NM_000094.3(COL7A1):c.6724G>A (p.Gly2242Arg)	CAGRGGTCTCCAGGTTTGCCTGG	
121912838	COL7A1	NM_000094.3(COL7A1):c.6091G>A (p.Gly2031Ser)	TCCRGCCTTGCCGGGGAGCCTGG	Recessive dystrophic epidermolysis bullosa
121912839	COL7A1	NM_000094.3(COL7A1):c.6859G>A (p.Gly2287Arg)	GTCRGACCTAAAGGAGAACCTGG	Recessive dystrophic epidermolysis bullosa, Nail disorder, nonsyndromic congenital, 8
121912842	COL7A1	NM_000094.3(COL7A1):c.6017G>A (p.Gly2006Asp)	AAGGRCGACCGTGGAGACCCTGG	Generalized dominant dystrophic epidermolysis bullosa
121912843	COL7A1	NM_000094.3(COL7A1):c.6044G>A (p.Gly2015Glu)	CCCTCAGGRGCCACCTGGTCTGG	Generalized dominant dystrophic epidermolysis bullosa
121912844	COL7A1	NM_000094.3(COL7A1):c.6100G>A (p.Gly2034Arg)	GCCRGGGAGCCTGGAAAGCCTGG	Generalized dominant dystrophic epidermolysis bullosa
121912846	COL7A1	NM_000094.3(COL7A1):c.6110G>A (p.Gly2037Glu)	CTGRAAAGCCTGGTATTCCCGGG, CCTGRAAAGCCTGGTATTCCCGG	Generalized dominant dystrophic epidermolysis bullosa
121912850	COL7A1	NM_000094.3(COL7A1):c.6227G>A (p.Gly2076Asp)	GATGRCCCTCCTGGACTCCCTGG	Generalized dominant dystrophic epidermolysis bullosa
121912851	COL7A1	NM_000094.3(COL7A1):c.7957G>A (p.Gly2653Arg)	CCCRGGCTGGCAGGACACAAAGG	Recessive dystrophic epidermolysis bullosa
387906605	COL7A1	NM_000094.3(COL7A1):c.4565G>A (p.Gly1522Glu)	CAGGRGCCACCAGGACCCACTGG	Transient bullous dermolysis of the newborn
121912847	COL7A1	NM_000094.3(COL7A1):c.4888C>T (p.Arg1630Ter)	CCCAGGACCTGTTGGCCCCYGAG, CCAGGACCTGTTGGCCCCYGAGG	
121912931	COL9A1	NM_001851.4(COL9A1):c.883C>T (p.Arg295Ter)	CCTACCCCTCCAGGGTGACYGAG, CCCCTCCAGGGTGACYGAGGTCC, CCCTCCAGGGTGACYGAGGTCCT	Stickler syndrome, type 4
387907076	COLEC11	NM_024027.4(COLEC11):c.610G>A (p.Gly204Ser)	CTTCATCRGCATCAACGACCTGG	Carnevale syndrome
312262904	COMP	NM_000095.2(COMP):c.2155G>A (p.Gly719Ser)	CATGCGGRGTGGCCGCCTGGGGG, CCATGCGGRGTGGCCGCCTGGGG	Pseudoachondroplastic spondyloepiphyseal dysplasia syndrome
137852655	COMP	NM_000095.2(COMP):c.2156G>A (p.Gly719Asp)	CATGCGGGRTGGCCGCCTGGGGG	Pseudoachondroplastic spondyloepiphyseal dysplasia syndrome

121918231	COQ2	NM_015697.7(COQ2):c.590G>A (p.Arg197His)	GCCAATCRTCCAATAGCCGCTGG	Coenzyme Q10 deficiency, primary 1
143441644	COQ4	NM_016035.4(COQ4):c.718C>T (p.Arg240Cys)	CCTGTACTATGAGCGGYGCTGGG	COENZYME Q10 DEFICIENCY, PRIMARY, 7
397514755	CORO1A	NM_001193333.2(CORO1A):c.400G>A (p.Val134Met)	AGCGTRTGGGCATTGTGGCCTGG	Immunodeficiency 8
28939711	COX15	NM_004376.5(COX15):c.649C>T (p.Arg217Trp)	CCCATGACATCCCTYGGGTCAGT	Congenital myasthenic syndrome, acetazolamide-responsive, Cardioencephalomyopathy, fatal infantile, due to cytochrome c oxidase deficiency 2
61733458	CP	NM_000096.3(CP):c.1652C>T (p.Thr551Ile)	CCARTGAATATATCTTTAGTGGG, CCCARTGAATATATCTTTAGTGG	Deficiency of ferroxidase, not specified
386134156	CP	NM_000096.3(CP):c.2701C>T (p.Arg901Ter)	CCCCCTGATTGTTTGTYGAAGAC, CCCCTGATTGTTTGTYGAAGACC, CCCTGATTGTTTGTYGAAGACCT	Deficiency of ferroxidase
114402678	CPA6	NM_020361.4(CPA6):c.809C>T (p.Ala270Val)	TTCTATTGRCATCCACTCCACGG	Febrile seizures, familial, 11, not provided
121917866	CPOX	NM_000097.5(CPOX):c.991C>T (p.Arg331Trp)	CCCATCGTGGAGAAYGGCGGGGC	Coproporphyria
121917871	CPOX	NM_000097.5(CPOX):c.85C>T (p.Gln29Ter)	CCCCGCGCCTGGTCCYAGTGCGG, CCCGCGCCTGGTCCYAGTGCGGC	Coproporphyria
28936374	CPT1A	NM_001876.3(CPT1A):c.2126G>A (p.Gly709Glu)	CAGCGRAGGGGGCTTTGGACCGG	Carnitine palmitoyltransferase I deficiency
80356780	CPT1A	NM_001876.3(CPT1A):c.2129G>A (p.Gly710Glu)	CAGCGGAGRGGGCTTTGGACCGG	Carnitine palmitoyltransferase I deficiency
80356794	CPT1A	NM_001876.3(CPT1A):c.1425G>A (p.Trp475Ter)	CTGRGCAGATGCGCCGATCGTGG	Carnitine palmitoyltransferase I deficiency
80356775	CPT1A	NM_001876.3(CPT1A):c.367C>T (p.Arg123Cys)	CCCTCATCGTCACCATGYGCTAC, CCTCATCGTCACCATGYGCTACT	Carnitine palmitoyltransferase I deficiency, not provided
80356779	CPT1A	NM_001876.3(CPT1A):c.1436C>T (p.Pro479Leu)	CCTGGGCAGATGCGCYGATCGTG	Carnitine palmitoyltransferase I deficiency, not provided
727503887	CPT2	NM_000098.2(CPT2):c.886C>T (p.Arg296Ter)	CCTGACCAGTGAGAACYGAGACA	CARNITINE PALMITOYLTRANSFERASE II DEFICIENCY, LATE-ONSET, Carnitine palmitoyltransferase II deficiency, infantile
74315296	CPT2	NM_000098.2(CPT2):c.1507C>T (p.Arg503Cys)	CCGCACTGAGACCATCYGCCCGG	
28939720	CRB1	NM_201253.2(CRB1):c.2234C>T (p.Thr745Met)	CCATGTTTGTCCGAAYGCTTCAA	Retinitis pigmentosa 12, not provided
730880377	CRB2	NM_173689.6(CRB2):c.189	CCGTTGCGACTGTGCCYGGCCCC	Ventriculomegaly with cystic kidney disease

		7C>T (p.Arg633Trp)		
587783476	CREBBP	NM_004380.2(CREBBP):c.286C>T (p.Gln96Ter)	CCAGCAGCCCCGTGYAGCAGGGC	Rubinstein-Taybi syndrome
587783475	CREBBP	NM_004380.2(CREBBP):c.2791C>T (p.Gln931Ter)	CCAGGTGACCCCGCAGCCTYAAA	Rubinstein-Taybi syndrome
587783479	CREBBP	NM_004380.2(CREBBP):c.3310C>T (p.Gln1104Ter)	CCTAGAAGCACTGTATCGAYAGG	Rubinstein-Taybi syndrome
587783490	CREBBP	NM_004380.2(CREBBP):c.4078C>T (p.Arg1360Ter)	CCGGGGAGGTTTTTGTCYGAGTG	Rubinstein-Taybi syndrome
587783509	CREBBP	NM_004380.2(CREBBP):c.598C>T (p.Gln200Ter)	CCATAGCTTAATTAATYAGGCTT	Rubinstein-Taybi syndrome
587783510	CREBBP	NM_004380.2(CREBBP):c.6088C>T (p.Gln2030Ter)	CCCCTTCCCCAGCAGYAGCCCAT, CCCTTCCCCAGCAGYAGCCCATG	Rubinstein-Taybi syndrome
137853932	CRLF1	NM_004750.4(CRLF1):c.397+1G>A	TTGGCCRTAAGTTGGCACCCAGG	Cold-induced sweating syndrome 1
137853926	CRLF1	NM_004750.4(CRLF1):c.538C>T (p.Gln180Ter)	CCCCGCAGGTGGTATGGCYAGGA, CCCGCAGGTGGTATGGCYAGGAC, CCGCAGGTGGTATGGCYAGGACA	Cold-induced sweating syndrome 1
137853930	CRLF1	NM_004750.4(CRLF1):c.413C>T (p.Pro138Leu)	CCAGTGCCCCCAGAGAAACYCGT	Cold-induced sweating syndrome 1
72659357	CRTAP	NM_006371.4(CRTAP):c.3G>A (p.Met1Ile)	GATRGAGCCGGGGCGCCGGGGG, CGATRGAGCCGGGGCGCCGGGG, GCGATRGAGCCGGGGCGCCGGG, CGCGATRGAGCCGGGGCGCCGG, GCGCGATRGAGCCGGGGCGCCG	Osteogenesis imperfecta type 7
104894672	CRX	NM_000554.4(CRX):c.121C>T (p.Arg41Trp)	CCCCCAGGAAGCAGCGGYGGGAG, CCCCAGGAAGCAGCGGYGGGAGC, CCCAGGAAGCAGCGGYGGGAGCG, CCAGGAAGCAGCGGYGGGAGCG	Cone-rod dystrophy 2, not provided

			C	
104894673	CRX	NM_000554.4(CRX):c.268C>T (p.Arg90Trp)	CCAGGTTTGGTTCAAGAACYGGA	Leber congenital amaurosis 7, not provided
74315440	CRYAA	NM_000394.3(CRYAA):c.27G>A (p.Trp9Ter)	CCTGRTTCAAGCGCACCCTGGGG, CCCTGRTTCAAGCGCACCCTGGG, CCCCTGRTTCAAGCGCACCCTGG	
397515623	CRYAA	NM_000394.3(CRYAA):c.160C>T (p.Arg54Cys)	CCGCCAGTCCCTCTTCYGCACCG	Cataract, autosomal dominant
397515624	CRYAA	NM_000394.3(CRYAA):c.34C>T (p.Arg12Cys)	CCAGCACCCCTGGTTCAAGYGCA	Cataract, autosomal dominant, multiple types, with microcornea
74315439	CRYAA	NM_000394.3(CRYAA):c.346C>T (p.Arg116Cys)	CCCGTGAGTTCCACYGCCGCTAC	Cataract, autosomal dominant
74315441	CRYAA	NM_000394.3(CRYAA):c.145C>T (p.Arg49Cys)	CCATCAGCCCCTACTACYGCCAG	Cataract, autosomal dominant
387907338	CRYAB	NM_001885.2(CRYAB):c.166C>T (p.Arg56Trp)	CCACCCTCCTTCCTGYGGGCACC	Posterior polar cataract type 2
397515555	CSF1R	NM_005211.3(CSF1R):c.1958G>A (p.Cys653Tyr)	GCCTRTACCCATGGAGGTAAGGG, AGCCTRTACCCATGGAGGTAAGG	Hereditary diffuse leukoencephalopathy with spheroids
690016564	CSF1R	NM_005211.3(CSF1R):c.2350G>A (p.Val784Met)	CGTAACRTGCTGTTGACCAATGG	Hereditary diffuse leukoencephalopathy with spheroids
281860268	CSF1R	NM_005211.3(CSF1R):c.1766G>A (p.Gly589Glu)	CCCTCGRAGCTGGAGCCTTTGGG, ACCCTCGRAGCTGGAGCCTTTGG	Hereditary diffuse leukoencephalopathy with spheroids
397515556	CSF1R	NM_005211.3(CSF1R):c.2329C>T (p.Arg777Trp)	CCCTCAGTGCATCCACYGGGACG, CCTCAGTGCATCCACYGGGACGT	Hereditary diffuse leukoencephalopathy with spheroids
690016562	CSF1R	NM_005211.3:c.2467C>T	CCTGTGAAGTGGATGGYCCCAGA	Hereditary diffuse leukoencephalopathy with spheroids
606231473	CSF3R	NM_000760.3(CSF3R):c.922C>T (p.Arg308Cys)	CCTACACCCTGCAGATAYGCTGC	Severe congenital neutropenia
796065343	CSF3R	NM_156039.3(CSF3R):c.1853C>T (p.Thr618Ile)	CCTCACCCTGATGAYCTTGACCC	Early T cell progenitor acute lymphoblastic leukemia
1064039	CST3	NM_000099.3(CST3):c.73G>A (p.Ala25Thr)	CCCGCGRCCGGCTCCAGTCCCGG	Age-related macular degeneration 11
545986367	CSTB	NM_000100.3(CSTB):c.136C>T (p.Gln46Ter)	CACCTRGCTCTTGAATGACACGG	not provided
387907080	CTC1	NM_025099.5(CTC1):c.775G>A (p.Val259Met)	CCACRTGTCCATCATCGTGCAGG	Cerebroretinal microangiopathy with calcifications and cysts
121913413	CTNNB1	NM_001904.3(CTNNB1):c.122C>T (p.Thr41Ile)	CCATTCTGGTGCCACTAYCACAG	Pilomatrixoma

35086888	CTNS	NM_001031681.2(CTNS):c.124G>A (p.Val42Ile)	CCGCAGGGTGAGGCTGAYGTTGG	Cystinosis, atypical nephropathic
515726209	CTRC	NM_007272.2(CTRC):c.217G>A (p.Ala73Thr)	TCACTGCCRCCCACTGCATCAGG	Hereditary pancreatitis
121909293	CTRC	NM_007272.2(CTRC):c.760C>T (p.Arg254Trp)	CCGGTAGTCTACACCYGGGTGTC	Hereditary pancreatitis, Pancreatitis, chronic, susceptibility to
104894209	CTSC	NM_001814.4(CTSC):c.856C>T (p.Gln286Ter)	CCCCAATCCTAAGCCCTYAGGAG, CCCAATCCTAAGCCCTYAGGAGG, CCAATCCTAAGCCCTYAGGAGGT	Papillon-Lef\xc3\xa8vre syndrome
587779409	CTSD	NM_001909.4(CTSD):c.470C>T (p.Ser157Leu)	CCAGGACACTGTGTYGGTGAGTC	Ceroid lipofuscinosis neuronal 10
3732378	CX3CR1	NM_001171174.1(CX3CR1):c.935C>T (p.Thr312Met)	AACCRTCTCAGTCACACTGAGGG, CAACCRTCTCAGTCACACTGAGG	Age-related macular degeneration 12
3732379	CX3CR1	NM_001171174.1(CX3CR1):c.841G>A (p.Val281Ile)	CCAGGAAAATCATAAYGTTGTAG	Age-related macular degeneration 12
104893624	CXCR4	NM_003467.2(CXCR4):c.1000C>T (p.Arg334Ter)	CCTCTCCAAAGGAAAGYGAGGTG	Warts, hypogammaglobulinemia, infections, and myelokathexis
121965009	CYB5R3	NM_000398.6(CYB5R3):c.316G>A (p.Val106Met)	CTTCRTGGACCTGGTCATCAAGG	METHEMOGLOBINEMIA, TYPE I
61732609	CYB5R3	NM_000398.6(CYB5R3):c.478C>T (p.Arg160Ter)	GTCRGATGGCGAACTTCCCTGGG, GGTCRGATGGCGAACTTCCCTGG	Methemoglobinemia type 2
121965014	CYB5R3	NM_000398.6(CYB5R3):c.229C>T (p.Gln77Ter)	CCCGTTCTGTCCTGCAGGCYAGC, CCGTTCTGTCCTGCAGGCYAGCA	Methemoglobinemia type 2
200872504	CYB5R3	NM_000398.6(CYB5R3):c.463+8G>C	CCAAGGGATTCCGACCYGAATCA	Methemoglobinemia type 2
137854590	CYBB	NM_000397.3(CYBB):c.466G>A (p.Ala156Thr)	TTTTRCTCGAAAGAGAATAAAGG	Granulomatous disease, chronic, X-linked, variant, not provided
104894139	CYP17A1	NM_000102.3(CYP17A1):c.1073G>A (p.Arg358Gln)	TCCRAGAGGTGCTTCGCCTCAGG	Isolated 17,20-lyase deficiency
104894153	CYP17A1	NM_000102.3(CYP17A1):c.287G>A (p.Arg96Gln)	GGGCRGCCTCAAATGGTAAGTGG	Complete combined 17-alpha-hydroxylase/17,20-lyase deficiency
104894154	CYP17A1	NM_000102.3(CYP17A1):c.374G>A (p.Arg125Gln)	GCTGCATCRAAGGCTGGCGATGG	Complete combined 17-alpha-hydroxylase/17,20-lyase deficiency
104894155	CYP17A1	NM_000102.3(CYP17A1):c.1247G>A (p.Arg416His)	AGCRTTTCTTGAATCCAGCGGGG, GAGCRTTTCTTGAATCCAGCGGG, AGAGCRTTTCTTGAATCCAGCGG	Complete combined 17-alpha-hydroxylase/17,20-lyase deficiency
104894142	CYP17A1	NM_000102.3(CYP17A1):c	CCATCCGAGAGGTGCTTYGCCTC	Complete combined 17-alpha-hydroxylase/17,20-lyase

		.1084C>T (p.Arg362Cys)		deficiency
104894145	CYP17A1	NM_000102.3(CYP17A1):c.1283C>T (p.Pro428Leu)	CCCAGCTCATCTCACYGTCAGTA, CCAGCTCATCTCACYGTCAGTAA	Complete combined 17-alpha-hydroxylase/17,20-lyase deficiency
104894149	CYP17A1	NM_000102.3(CYP17A1):c.1039C>T (p.Arg347Cys)	CCAACTATCAGTGACYGTAACCG	Combined partial 17-alpha-hydroxylase/17,20-lyase deficiency
28936700	CYP1B1	NM_000104.3(CYP1B1):c.182G>A (p.Gly61Glu)	ATCGRAAACGCGGCGGCGGTGGG, GATCGRAAACGCGGCGGCGGTGG, ACTGATCGRAAACGCGGCGGCGG	Glaucoma, congenital
201824781	CYP1B1	NM_000104.3(CYP1B1):c.155C>T (p.Pro52Leu)	ACGGGCCCRGGGGCGCGGACCGG	Glaucoma, primary open angle, juvenile-onset
79204362	CYP1B1	NM_000104.3(CYP1B1):c.1103G>A (p.Arg368His)	CCCATACAAGGCAGAYGGTCCCT, CCATACAAGGCAGAYGGTCCCTC	Glaucoma, congenital, Coloboma, not provided
151344503	CYP21A2	NM_000500.7(CYP21A2):c.1217G>A (p.Trp406Ter)	GTTCTRGCCTGGTATGTGGGGGG, AGTTCTRGCCTGGTATGTGGGGG, GAGTTCTRGCCTGGTATGTGGGG, TGAGTTCTRGCCTGGTATGTGGG	21-hydroxylase deficiency
7769409	CYP21A2	NM_000500.7(CYP21A2):c.1069C>T (p.Arg357Trp)	CCGAGGTGCTGCGCCTGYGGCCC	21-hydroxylase deficiency
6445	CYP21A2	NM_000500.7(CYP21A2):c.1360C>T (p.Pro454Ser)	CCTTCACGCTGCTGYCCTCCGGG	21-hydroxylase deficiency
9378251	CYP21A2	NM_000500.7(CYP21A2):c.92C>T (p.Pro31Leu)	CCGGAGCCTCCACCTCCYGCCTC	21-hydroxylase deficiency
7755898	CYP21A2	NM_000500.7(CYP21A2):c.955C>T (p.Gln319Ter)	CCAGATTCAGCAGCGACTGYAGG	21-hydroxylase deficiency
387907324	CYP24A1	NM_000782.4(CYP24A1):c.964G>A (p.Glu322Lys)	CACARAGCTCCAGCTGGCTGCGG	Idiopathic hypercalcemia of infancy
121908097	CYP27A1	NM_000784.3(CYP27A1):c.1421G>A (p.Arg474Gln)	TCCRGGCCTGCCTGGGCCGCAGG	Cholestanol storage disease
121908099	CYP27A1	NM_000784.3(CYP27A1):c.1214G>A (p.Arg405Gln)	AAACTCCCRGATCATAGAAAAGG	Cholestanol storage disease
587778797	CYP27A1	NM_000784.3(CYP27A1):c.446+1G>A	CACRTGAGCTGGGGCCTGAAGGG, CCACRTGAGCTGGGGCCTGAAGG	Cholestanol storage disease
376230356	CYP27A1	NM_000784.3(CYP27A1):c.380G>A (p.Arg127Gln)	TACRGAACGACATGGAGCTATGG	Cholestanol storage disease
72551314	CYP27A1	NM_000784.3(CYP27A1):c	CCACTGGTACCAGCTGCGCYAGG	Cholestanol storage disease

		.475C>T (p.Gln159Ter)		
72551316	CYP27A1	NM_000784.3(CYP27A1):c.745C>T (p.Gln249Ter)	CCATCGGGTTAATGTTCYAGAAC	Cholestanol storage disease
121908098	CYP27A1	NM_000784.3(CYP27A1):c.1420C>T (p.Arg474Trp)	CCCTTTGGCTATGGGGTCYGGGC, CCTTTGGCTATGGGGTCYGGGCC	Cholestanol storage disease
587778787	CYP27A1	NM_000784.3(CYP27A1):c.1402C>T (p.Pro468Ser)	CCCATTTGGCTCTGTGYCCTTTG, CCATTTGGCTCTGTGYCCTTTGG	Cholestanol storage disease
28934604	CYP27B1	NM_000785.3(CYP27B1):c.320G>A (p.Arg107His)	AGCRCTGCAGCTTCTCGCCCTGG	Vitamin D-dependent rickets, type 1
118204008	CYP27B1	NM_000785.3(CYP27B1):c.1226C>T (p.Thr409Ile)	CCAGACGCTGGTCAYTCTGTGTC	Vitamin D-dependent rickets, type 1
118204011	CYP27B1	NM_000785.3(CYP27B1):c.1027C>T (p.Leu343Phe)	CCCCGAAGTCCAGACAGCAYTCC, CCCGAAGTCCAGACAGCAYTCCA, CCGAAGTCCAGACAGCAYTCCAC	Vitamin D-dependent rickets, type 1
199476187	CYP4V2	NM_207352.3(CYP4V2):c.283G>A (p.Gly95Arg)	GGTCRGGCCAGTGCCCATGGTGG, CTGGGTCRGGCCAGTGCCCATGG	Bietti crystalline corneoretinal dystrophy
199476198	CYP4V2	NM_207352.3(CYP4V2):c.1020G>A (p.Trp340Ter)	AACTGRTCCTTATACCTGTTGGG, AAACTGRTCCTTATACCTGTTGG	Bietti crystalline corneoretinal dystrophy
119103284	CYP4V2	NM_207352.3(CYP4V2):c.1523G>A (p.Arg508His)	TTCRTCCAAGTAATGGCATCTGG	Bietti crystalline corneoretinal dystrophy
199476202	CYP4V2	NM_207352.3(CYP4V2):c.1187C>T (p.Pro396Leu)	CCTTTTTCCTTCTGTTCYTTTAT	Bietti crystalline corneoretinal dystrophy
121908611	CYP7B1	NM_004820.3(CYP7B1):c.1250G>A (p.Arg417His)	TATGATCRTTTTATAGAAGATGG	Spastic paraplegia 5A
72554620	CYP7B1	NM_004820.3(CYP7B1):c.1162C>T (p.Arg388Ter)	CCGGGGACTACTGTGTGYGAAAG	Spastic paraplegia 5A, Bile acid synthesis defect, congenital, 3
193922955	DAG1	NM_001165928.3(DAG1):c.575C>T (p.Thr192Met)	CCTGTGACTGTTTTGAYGGTGAT	Limb-girdle muscular dystrophy-dystroglycanopathy, type C9, not provided
121918212	DARS2	NM_018122.4(DARS2):c.1837C>T (p.Leu613Phe)	CCTTCCGGGGACATGACYTCATG	Leukoencephalopathy with Brainstem and Spinal Cord Involvement and Lactate Elevation
200670286	DARS2	NM_018122.4(DARS2):c.1825C>T (p.Arg609Trp)	CCTTCCCAAAGTCCTTCYGGGGA	Leukoencephalopathy with Brainstem and Spinal Cord Involvement and Lactate Elevation
267606760	DBH	NM_000787.3(DBH):c.301G>A (p.Val101Met)	CTCRTGGTGCTCTGGACCGATGG	Dopamine beta hydroxylase deficiency
185492864	DBT	NM_001918.3(DBT):c.901C>T (p.Arg301Cys)	ATTCCACRAGCAAATGCAATGGG,	Maple syrup urine disease, not provided

			AATTCCACRAGCAAATGCAATGG	
72466485	DCTN1	NM_004082.4(DCTN1):c.211G>A (p.Gly71Arg)	ATGATRGAACTGTTCAAGGCAGG	Perry syndrome
104894779	DCX	NM_178152.2(DCX):c.184G>A (p.Asp62Asn)	ATGGGRACCGCTACTTCAAGGGG, AATGGGRACCGCTACTTCAAGGG, CAATGGGRACCGCTACTTCAAGG	Lissencephaly, X-linked, Subcortical laminar heterotopia, X-linked
587783544	DCX	NM_178151.2(DCX):c.364G>A (p.Gly122Arg)	AGGAARGTAATTTAAATAGTGGG, GAGGAARGTAATTTAAATAGTGG	Heterotopia
587783527	DCX	NM_178151.2(DCX):c.182G>A (p.Gly61Glu)	ATGRGGACCGCTACTTCAAGGGG, AATGRGGACCGCTACTTCAAGGG, CAATGRGGACCGCTACTTCAAGG	Heterotopia
587783589	DCX	NM_178151.2(DCX):c.809-1G>A	CCARAATGCCGAGTCATGAAGGG, CCCARAATGCCGAGTCATGAAGG	Heterotopia
104894780	DCX	NM_178151.2(DCX):c.574C>T (p.Arg192Trp)	CCGCAGTGGGGTGAAGCCTYGGA	Lissencephaly, X-linked, Subcortical laminar heterotopia, X-linked, Heterotopia
587783519	DCX	NM_178151.2(DCX):c.115C>T (p.Arg39Ter)	CCCACTGTAGCTTCTACYGAACC, CCACTGTAGCTTCTACYGAACCA	Heterotopia
587783522	DCX	NM_178151.2(DCX):c.130C>T (p.Gln44Ter)	CCGAACCAGAACCTTGYAGGCAC	Heterotopia
587783535	DCX	NM_178151.2(DCX):c.232C>T (p.Arg78Cys)	CCTCTGACCGTTTTYGCAGCTTT	Heterotopia
587783541	DCX	NM_178151.2(DCX):c.304C>T (p.Arg102Cys)	CCTGCCTCAGGGAGTGYGTTACA	Heterotopia
587783554	DCX	NM_178151.2(DCX):c.478C>T (p.Gln160Ter)	CCAATATGAAAGCCCCCYAGTCC	Heterotopia
587783590	DCX	NM_178151.2(DCX):c.814C>T (p.Arg272Ter)	CCTTTTGCCCCAGAATGCYGAGT	Heterotopia
587783592	DCX	NM_178151.2(DCX):c.907C>T (p.Arg303Ter)	CCCTGGTCCTATGCGCYGAAGCA, CCTGGTCCTATGCGCYGAAGCAA	Heterotopia
121434641	DDB2	NM_000107.2(DDB2):c.937C>T (p.Arg313Ter)	CCAGAAGAGCGAGATCYGAGTTT	Xeroderma pigmentosum, group E
137853208	DDC	NM_001082971.1(DDC):c.749C>T (p.Ser250Phe)	CCACAACATGCTGCTYCTTTGAC	Deficiency of aromatic-L-amino-acid decarboxylase

121964863	DDR2	NM_001014796.1(DDR2):c.2254C>T (p.Arg752Cys)	CCGGGCAGTGCTCCCTATCYGCT	Spondylometaepiphyseal dysplasia short limb-hand type
796052231	DDX3X	NM_001356.4(DDX3X):c.1126C>T (p.Arg376Cys)	CCTCCAAAGGGTGTCYGCCACAC	not provided, MENTAL RETARDATION, X-LINKED 102
796052234	DDX3X	NM_001356.4(DDX3X):c.1462C>T (p.Arg488Cys)	CCCTTCACCAGTTCYGCTCAGGA	not provided
796052236	DDX3X	NM_001356.4(DDX3X):c.1490C>T (p.Ala497Val)	CCCAATTTTAGTGGYTACAGCAG	not provided
587777408	DEAF1	NM_021008.3(DEAF1):c.670C>T (p.Arg224Trp)	CCCCRGCCGCCTGCAAGGAAGGG, TCCCCRGCCGCCTGCAAGGAAGG	Mental retardation, autosomal dominant 24
587777458	DEPDC5	NM_001242896.1(DEPDC5):c.3259C>T (p.Arg1087Ter)	CCTACATGGACAGCCCAYGAAAG	Epilepsy, partial, with variable foci
748323823	DES	NM_001927.3(DES):c.1371+1G>A	GGGAGRTAAGTGGTCTGTCTGGG, GGGGAGRTAAGTGGTCTGTCTGG	not provided
57694264	DES	NM_001927.3(DES):c.1201G>A (p.Glu401Lys)	ATGTGRAGATTGCCACCTACCGG	not provided
61726467	DES	NM_001927.3(DES):c.1237G>A (p.Glu413Lys)	GGGAGAGRAGAGCCGGTGAGGGG, AGGGAGAGRAGAGCCGGTGAGGG	not provided
62635763	DES	NM_001927.3(DES):c.1255C>T (p.Pro419Ser)	CCCTTTTAGGATCAATCTCYCCA, CCTTTTAGGATCAATCTCYCCAT	Myofibrillar myopathy 1, not provided
62636495	DES	NM_001927.3(DES):c.38C>T (p.Ser13Phe)	CCAGCCAGCGCGTGTCCTYCTAC, CCAGCGCGTGTCCTYCTACCGCC	Dilated cardiomyopathy 1I, Myofibrillar myopathy 1, not provided
397517255	DFNB31	NM_015404.3(DFNB31):c.1267C>T (p.Arg423Ter)	CCTGGGGAACCAGACAYGAGTGC	Usher syndrome, type 2D
779760634	DFNB31	NM_015404.3(DFNB31):c.1417-1G>A	CCTCAGAGAGGAGTGAGAAYTGG	Deafness, autosomal recessive 31
137852839	DFNB31	NM_015404.3(DFNB31):c.2332C>T (p.Arg778Ter)	CCAGCGCCCCAGGCYGAGGAAGG	Deafness, autosomal recessive 31
398123008	DGKE	NM_003647.2(DGKE):c.127C>T (p.Gln43Ter)	CCTTCTGGTGTAGCCTCYAGCGG	Nephrotic syndrome, type 7
104893631	DGUOK	NM_080916.2(DGUOK):c.425G>A (p.Arg142Lys)	AGARGTCTGTGTACAGTGACAGG	Mitochondrial DNA-depletion syndrome 3, hepatocerebral
121909764	DHCR7	NM_001360.2(DHCR7):c.730G>A (p.Gly244Arg)	GGCGCCCCRGGATCGTCGCCTGG	Smith-Lemli-Opitz syndrome

80338857	DHCR7	NM_001360.2(DHCR7):c.725G>A (p.Arg242His)	GGCRCCCCGGGATCGTCGCCTGG	Smith-Lemli-Opitz syndrome
398123607	DHCR7	NM_001360.2(DHCR7):c.841G>A (p.Val281Met)	TCTACRTGATTGACTTCTTCTGG	Smith-Lemli-Opitz syndrome, not provided
80338853	DHCR7	NM_001360.2(DHCR7):c.278C>T (p.Thr93Met)	CCAAGACTCCACCTATAAYGAGG	Smith-Lemli-Opitz syndrome, not provided
104886035	DHCR7	NM_001360.2(DHCR7):c.151C>T (p.Pro51Ser)	CCTACTGCTGTTCGCCYCCTTCA	Smith-Lemli-Opitz syndrome, not provided
267606766	DHODH	NM_001361.4(DHODH):c.454G>A (p.Gly152Arg)	CAGTCACRGGCTTTCAGTGGTGG	Miller syndrome
201230446	DHODH	NM_001361.4(DHODH):c.403C>T (p.Arg135Cys)	CCCTAGACCCAGAGTCTTCYGCC, CCTAGACCCAGAGTCTTCYGCCT	Miller syndrome
199422247	DKC1	NM_001363.4(DKC1):c.911G>A (p.Ser304Asn)	AAGACARTGCAGTAAGTTCCGGG, AAAGACARTGCAGTAAGTTCCGG	Dyskeratosis congenita X-linked
121912289	DKC1	NM_001363.4(DKC1):c.1226C>T (p.Pro409Leu)	CCCACAGACAGCACACYTGCCAC, CCACAGACAGCACACYTGCCACC	Dyskeratosis congenita X-linked
121964992	DLD	NM_000108.4(DLD):c.1123G>A (p.Glu375Lys)	TGTGTTRAAGGAATGGCTGGTGG	Maple syrup urine disease, type 3, not provided
796065346	DLL4	NM_019074.3(DLL4):c.1169G>A (p.Cys390Tyr)	GAATRTCCCCCCAACTTCACCGG	Adams-Oliver syndrome, ADAMS-OLIVER SYNDROME 6
398122853	DMD	NM_004006.2(DMD):c.9G>A (p.Trp3Ter)	GCTTTGRTGGGAAGAAGTAGAGG	Duchenne muscular dystrophy, Becker muscular dystrophy, Dilated cardiomyopathy 3B, not provided
398123936	DMD	NM_004006.2(DMD):c.336G>A (p.Trp112Ter)	TTTGRAATATAATCCTCCACTGG	Duchenne muscular dystrophy, Becker muscular dystrophy, Dilated cardiomyopathy 3B
398123939	DMD	NM_004006.2(DMD):c.3433-1G>A	GAARGTCTATGCCAGAAAGGAGG, GTGGAARGTCTATGCCAGAAAGG	Dilated cardiomyopathy 3B
398124032	DMD	NM_004006.2(DMD):c.649+1G>A	CCTGAAGRTTGGTAAATTTCTGG	Duchenne muscular dystrophy, Becker muscular dystrophy, Dilated cardiomyopathy 3B
398124076	DMD	NM_004006.2(DMD):c.8668G>A (p.Glu2890Lys)	AGCCCAGARGTAATTGAATGTGG	Dilated cardiomyopathy 3B
398124096	DMD	NM_004006.2(DMD):c.9564-1G>A	ATAATARGGGACGAACAGGGAGG	Duchenne muscular dystrophy, Becker muscular dystrophy, Dilated cardiomyopathy 3B
794727499	DMD	NM_004006.2(DMD):c.133C>T (p.Gln45Ter)	CCTCTTCAGTGACCTAYAGGATG	Duchenne muscular dystrophy, Becker muscular dystrophy
128626233	DMD	NM_004006.2(DMD):c.178C>T (p.Gln60Ter)	CCTCGAAGGCCTGACAGGGYAAA	Duchenne muscular dystrophy

128626238	DMD	NM_000109.3(DMD):c.700C>T (p.Gln234Ter)	CCAAGTTTTGCCTCAAYAAGTGA	Duchenne muscular dystrophy
128626245	DMD	NM_004006.2(DMD):c.3121C>T (p.Gln1041Ter)	CCAGCTGGTTGAGCATTGTYAAA	Duchenne muscular dystrophy, Dilated cardiomyopathy 3B
398123865	DMD	NM_004006.2(DMD):c.1615C>T (p.Arg539Ter)	CCAGGTATTGGGAGATYGATGGG	Duchenne muscular dystrophy, Becker muscular dystrophy, Dilated cardiomyopathy 3B
398123903	DMD	NM_004006.2(DMD):c.2650C>T (p.Gln884Ter)	CCGGCTATCAGATCTTYAACCTC	Duchenne muscular dystrophy, Becker muscular dystrophy, Dilated cardiomyopathy 3B
398123912	DMD	NM_004006.2(DMD):c.2866C>T (p.Gln956Ter)	CCATCAGGACATGGGTCYAGCAG	Dilated cardiomyopathy 3B
398123954	DMD	NM_004006.2(DMD):c.4405C>T (p.Gln1469Ter)	CCAGCCAATTTTGAGYAGCGTCT	Dilated cardiomyopathy 3B
398123990	DMD	NM_004006.2(DMD):c.5353C>T (p.Gln1785Ter)	CCTTTGAAGGAATTGGAGYAGTT	Dilated cardiomyopathy 3B
398123999	DMD	NM_004006.2(DMD):c.583C>T (p.Arg195Ter)	CCAGCAGTCAGCCACACAAYGAC	Duchenne muscular dystrophy, Becker muscular dystrophy, Dilated cardiomyopathy 3B
398124058	DMD	NM_004006.2(DMD):c.7894C>T (p.Gln2632Ter)	CCAAAGACCTCCGCYAGTGGCAG	Duchenne muscular dystrophy, Becker muscular dystrophy, Dilated cardiomyopathy 3B
398124099	DMD	NM_004006.2(DMD):c.961-5831C>T	CCTTTGTGACCTTTGGYAAGTCA	Dilated cardiomyopathy 3B
368260932	DNAH11	NM_001277115.1(DNAH11):c.8698C>T (p.Arg2900Ter)	CCAATTTGTACATCYGAACTGGA	Ciliary dyskinesia, primary, 7
567050969	DNAH8	NM_001206927.1(DNAH8):c.2419C>T (p.Arg807Ter)	CCACGCTTTTTGTGYGACATCCA	Kartagener syndrome, not provided
730882139	DNAJB2	NM_001039550.1(DNAJB2):c.229+1G>A	GACAGRTAGGTGGAGTGGTGAGG	Charcot-Marie-Tooth disease, Spinal muscular atrophy, distal, autosomal recessive, 5
398122405	DNAJC6	NM_001256865.1(DNAJC6):c.2200C>T (p.Gln734Ter)	CCCCACTCCTCTCCCYAGAACCG, CCCACTCCTCTCCCYAGAACCGA	Parkinson disease 19, juvenile-onset
121909089	DNM2	NM_001005360.2(DNM2):c.1106G>A (p.Arg369Gln)	CGAGCRGTTCCCATTTGAGCTGG	Myopathy, centronuclear, 1, Myopathy, centronuclear
796065342	DNMT3A	NM_175629.2(DNMT3A):c.1204C>T (p.Gln402Ter)	CCAAGGCCGTGGAGGTGYAGAAC	Early T cell progenitor acute lymphoblastic leukemia
121908943	DNMT3B	NM_006892.3(DNMT3B):c.1807G>A (p.Ala603Thr)	CGTCRCTTCTGAAGTGTGTGAGG	Centromeric instability of chromosomes 1,9 and 16 and immunodeficiency
547940069	DNMT3B	NM_175850.2(DNMT3B):c.2397-11G>A	TCCRGTACCCCCAGGATCTTTGG	Centromeric instability of chromosomes 1,9 and 16 and immunodeficiency
780318765	DOCK2	NM_004946.2(DOCK2):c.3310C>T (p.Arg1104Trp)	CCCTGAGGCTGAGCTCYGGAAAG,	IMMUNODEFICIENCY 40

			CCTGAGGCTGAGCTCYGGAAAGC	
587777484	DOCK7	NM_033407.3(DOCK7):c.3616C>T (p.Arg1206Ter)	CCACTCRAGCCTTTATCTGAGGG, GCCACTCRAGCCTTTATCTGAGG	Epileptic encephalopathy, early infantile, 23
397515322	DPAGT1	NM_001382.3(DPAGT1):c.161+5G>A	TGARCAGCGGCACACGGGTCCGG, AGATGTGARCAGCGGCACACGGG	Congenital disorder of glycosylation type 1J
397515640	DSG1	NM_001942.3(DSG1):c.601C>T (p.Gln201Ter)	CCTTCAAGATTATAAGAYAAGAA	Keratosis palmoplantaris striata 1
751012696	DSG2	NM_001943.3(DSG2):c.889G>A (p.Asp297Asn)	GTGTTCRATGCAGATGAAATAGG	Cardiomyopathy
121913008	DSG2	NM_001943.3(DSG2):c.137G>A (p.Arg46Gln)	TAGTGCRGCAAAAGCGCGCCTGG	Arrhythmogenic right ventricular cardiomyopathy, type 10, Cardiomyopathy
121913013	DSG2	NM_001943.3(DSG2):c.166G>A (p.Val56Met)	GCCCCCRTGGCTCTTCGGGAGGG, CGCCCCCRTGGCTCTTCGGGAGG	Arrhythmogenic right ventricular cardiomyopathy, Arrhythmogenic right ventricular cardiomyopathy, type 10, Catecholaminergic polymorphic ventricular tachycardia, Dilated cardiomyopathy 1BB, Primary familial hypertrophic cardiomyopathy, Cardiomyopathy, not specified
201564919	DSG2	NM_001943.3(DSG2):c.1912G>A (p.Gly638Arg)	CATTGCRGAAAGGGCGCCAAAGG	Arrhythmogenic right ventricular cardiomyopathy, Cardiomyopathy, not specified
794728083	DSG2	NM_001943.3(DSG2):c.769C>T (p.Gln257Ter)	CCTGTAAAACAAGCTYAAGTTCA	Cardiomyopathy
121912998	DSP	NM_004415.2(DSP):c.88G>A (p.Val30Met)	GAGRTGACCAGCGGCGGCGGGGG, CGAGRTGACCAGCGGCGGCGGGG, ACGAGRTGACCAGCGGCGGCGGG, TACGAGRTGACCAGCGGCGGCGG	Arrhythmogenic right ventricular cardiomyopathy, Arrhythmogenic right ventricular cardiomyopathy, type 8, not specified, not provided
121912999	DSP	NM_004415.2(DSP):c.8501G>A (p.Arg2834His)	CTCRCTCCGGATCTCGCTCCGGG, TCTCRCTCCGGATCTCGCTCCGG	Arrhythmogenic right ventricular cardiomyopathy, type 8
397516943	DSP	NM_004415.2(DSP):c.478C>T (p.Arg160Ter)	CCATCAGTGTCCCTYGAGTCCGC	Arrhythmogenic right ventricular cardiomyopathy, type 8, not provided
767643821	DSP	NM_004415.2(DSP):c.3805C>T (p.Arg1269Ter)	CCACTGAGCAGCGAAGGYGAGCT	not provided
397509411	DYNC1H1	NM_001376.4(DYNC1H1):c.10151G>A (p.Arg3384Gln)	AATCRGGCTTCCCTGGCTTGCGG	Mental retardation, autosomal dominant 13
387906740	DYNC1H	NM_001376.4(DYNC1H1):	TGAARAGGATGCTCTCAGCTGGG	Mental retardation, autosomal dominant 13

	1	c.4552G>A (p.Glu1518Lys)	, TTGAARAGGATGCTCTCAGCTGG	
201479015	DYNC2H1	NM_001080463.1(DYNC2H1):c.11747G>A (p.Gly3916Asp)	TAGRTGCCAAAGATGTACAATGG	Short-rib thoracic dysplasia 3 with or without polydactyly, not provided
367643250	DYRK1B	NM_004714.2(DYRK1B):c.304C>T (p.Arg102Cys)	CACTGCRCACGATGTAGTCATGG	Abdominal obesity-metabolic syndrome 3
794727343	DYSF	NM_003494.3(DYSF):c.1956G>A (p.Trp652Ter)	CTGRGGTAACGTGAAACCTGTGG	Miyoshi muscular dystrophy 1, Limb-girdle muscular dystrophy, type 2B, Myopathy, distal, with anterior tibial onset
201869739	DYSF	NM_003494.3(DYSF):c.937+1G>A	CCCRTGAGTTCTCACCACTTTGG	Limb-girdle muscular dystrophy, type 2B, not provided
745891180	DYSF	NM_003494.3(DYSF):c.5057+5G>A	GTGTGTACRTGGATGGGGGCTGG	Limb-girdle muscular dystrophy, type 2B
398123763	DYSF	NM_003494.3(DYSF):c.1053+1G>A	GCCTRTGAGTACATTTCCCTGGG, CGCCTRTGAGTACATTTCCCTGG	Limb-girdle muscular dystrophy, type 2B, not provided
398123794	DYSF	NM_003494.3(DYSF):c.5509G>A (p.Asp1837Asn)	TGAGCRACATTTATGTGAAAGGG , ATGAGCRACATTTATGTGAAAGG	Limb-girdle muscular dystrophy, type 2B, not provided
794727636	DYSF	NM_003494.3(DYSF):c.265C>T (p.Arg89Ter)	CCAAGGTCCCACTCYGAGAGGTC	Miyoshi muscular dystrophy 1, Limb-girdle muscular dystrophy, type 2B
727503911	DYSF	NM_003494.3(DYSF):c.3832C>T (p.Gln1278Ter)	CCTCTTTTGAGCTCATCYAGAGA	Limb-girdle muscular dystrophy, type 2B, not provided
398123773	DYSF	NM_003494.3(DYSF):c.2311C>T (p.Gln771Ter)	CCCTGCAGCTCTGGAGYAGGCGG , CCTGCAGCTCTGGAGYAGGCGGA	Limb-girdle muscular dystrophy, type 2B, not provided
398123789	DYSF	NM_003494.3(DYSF):c.4756C>T (p.Arg1586Ter)	CCGTATCTACATTGTCYGAGCAT	Miyoshi muscular dystrophy 1, Limb-girdle muscular dystrophy, type 2B, not provided
397514594	EARS2	NM_001083614.1(EARS2):c.500G>A (p.Cys167Tyr)	TCGGTRCAGGAACATGAGCCAGG	Combined oxidative phosphorylation deficiency 12
376103091	EARS2	NM_001083614.1(EARS2):c.322C>T (p.Arg108Trp)	CCCCRGCGGGGGCTCTCATCAGG	Combined oxidative phosphorylation deficiency 12
104894792	EBP	NM_006579.2(EBP):c.386G>A (p.Trp129Ter)	TGTRGGGACCACTCAGCCTGTGG	Chondrodysplasia punctata 2 X-linked dominant
104894794	EBP	NM_006579.2(EBP):c.587G>A (p.Trp196Ter)	GCCCTGTRGCTGGTGCTGCCTGG	Chondrodysplasia punctata 2 X-linked dominant
104894798	EBP	NM_006579.2(EBP):c.87G>A (p.Trp29Ter)	CCCACCTGRCATATACTGGCTGG	Chondrodysplasia punctata 2 X-linked dominant
104894799	EBP	NM_006579.2(EBP):c.187C	CCCATTGGGGACTTGGCGGYGAC	Chondrodysplasia punctata 2 X-linked dominant

		>T (p.Arg63Ter)	, CCATTGGGGACTTGGCGGYGACT	
587783613	EBP	NM_006579.2(EBP):c.328C >T (p.Arg110Ter)	CCAAGGGAGACAGCYGATACAT C	Chondrodysplasia punctata 2 X-linked dominant
587776498	ECHS1	NM_004092.3(ECHS1):c.5 C>T (p.Ala2Val)	CCAGAGAGCCATGGYCGCCCTGC	Leigh disease, MITOCHONDRIAL SHORT-CHAIN ENOYL-CoA HYDRATASE 1 DEFICIENCY
121909114	ECM1	NM_004425.3(ECM1):c.10 36C>T (p.Gln346Ter)	CCAGCTGGAGAGGGAGTTCYAGC	Lipid proteinosis
121909115	ECM1	NM_004425.3(ECM1):c.15 7C>T (p.Arg53Ter)	CCCCTCCCCACCCCTATCCYGAA, CCCTCCCCACCCCTATCCYGAAG, CCTCCCCACCCCTATCCYGAAGC, CCCCACCCCTATCCYGAAGCCTC	Lipid proteinosis
397516677	EDA	NM_001399.4(EDA):c.871 G>A (p.Gly291Arg)	CAGCRGGGAGCTGGAGGTACTGG	Hypohidrotic X-linked ectodermal dysplasia
727504537	EDA	NM_001399.4(EDA):c.396 +1G>A	CCAGRTGAGTCACCTAGTAGGGG , ACCAGRTGAGTCACCTAGTAGGG , CACCAGRTGAGTCACCTAGTAGG	Hypohidrotic X-linked ectodermal dysplasia
132630310	EDA	NM_001399.4(EDA):c.67C >T (p.Gln23Ter)	CCGCGGGAGCGAGGGAGCYAGG G	Hypohidrotic X-linked ectodermal dysplasia
132630315	EDA	NM_001399.4(EDA):c.626 C>T (p.Pro209Leu)	CCAGGGATTCCTGGAATTCYAGG	Hypohidrotic X-linked ectodermal dysplasia, not specified
727503007	EDA	NM_001399.4(EDA):c.676 C>T (p.Gln226Ter)	CCAGGTCCTCCTGGTCCTYAAGG	Hypohidrotic X-linked ectodermal dysplasia
121908450	EDAR	NM_022336.3(EDAR):c.26 6G>A (p.Arg89His)	AGGCRTCACAAAGACTGTGAGGG , CAGGCRTCACAAAGACTGTGAGG	Autosomal recessive hypohidrotic ectodermal dysplasia syndrome
121908453	EDAR	NM_022336.3(EDAR):c.12 59G>A (p.Arg420Gln)	GATTGAGCRGCTGGATGCTGTGG	Autosomal dominant hypohidrotic ectodermal dysplasia
121908452	EDAR	NM_022336.3(EDAR):c.10 72C>T (p.Arg358Ter)	CCTCGAGAAGACTAGCYGAATGC	
74315309	EDARADD	NM_080738.3(EDARADD) :c.424G>A (p.Glu142Lys)	CTATGACRAATTGTGCTTCCTGG	Autosomal recessive hypohidrotic ectodermal dysplasia syndrome, Ectodermal dysplasia 11b, hypohidrotic/hair/tooth type, autosomal recessive
121434491	EFEMP1	NM_001039348.2(EFEMP1):c.1033C>T (p.Arg345Trp)	CCACAAATGAATGCYGGGAGGA T	Doyne honeycomb retinal dystrophy, Malattia leventinese
193302866	EFEMP2	NM_016938.4(EFEMP2):c. 800G>A (p.Cys267Tyr)	TTCTCCTRCCACTGCCCACAGGG, TTTCTCCTRCCACTGCCCACAGG	Autosomal recessive cutis laxa type IA, Autosomal recessive cutis laxa type 1B

119489102	EFEMP2	NM_016938.4(EFEMP2):c.835C>T (p.Arg279Cys)	CCAGCTGCTGGCCACAYGCCTCT	Autosomal recessive cutis laxa type IA, Autosomal recessive cutis laxa type 1B
387906878	EFTUD2	NM_004247.3(EFTUD2):c.2770C>T (p.Gln924Ter)	CCGCCCCTTGGAGCCAYAGCCAG	Growth and mental retardation, mandibulofacial dysostosis, microcephaly, and cleft palate
104894161	EGR2	NM_000399.3(EGR2):c.1075C>T (p.Arg359Trp)	CCGCTCTGACGAGCTGACAYGGC	Dejerine-Sottas disease, Charcot-Marie-Tooth disease, type ID
587777208	EIF2AK4	NM_001013703.3(EIF2AK4):c.3448C>T (p.Arg1150Ter)	CCGCGCAAGTTAGATYGATTTCA	Familial pulmonary capillary hemangiomatosis
104894427	EIF2B2	NM_014239.3(EIF2B2):c.547C>T (p.Arg183Ter)	CCTCAAAGAGGCTGCCYGAAAGA	Ovarioleukodystrophy
113994033	EIF2B4	NM_001034116.1(EIF2B4):c.1070G>A (p.Arg357Gln)	GCGGTTTCRGGTGGTAGTGGTGG	Leukoencephalopathy with vanishing white matter
113994037	EIF2B4	NM_001034116.1(EIF2B4):c.1191+1G>A	AGAGRTAAGTACAGAGGAAAAGG	Leukoencephalopathy with vanishing white matter
113994027	EIF2B4	NM_001034116.1(EIF2B4):c.683C>T (p.Ala228Val)	CCAATGCCCGGTGTATTGYCCTG	Leukoencephalopathy with vanishing white matter
113994055	EIF2B5	NM_003907.2(EIF2B5):c.583C>T (p.Arg195Cys)	CCCCCAGCCACCCAACTYGTTGC, CCCCAGCCACCCAACTYGTTGCC, CCCAGCCACCCAACTYGTTGCCA, CCAGCCACCCAACTYGTTGCCAC	Ovarioleukodystrophy
119484086	ELAC2	NM_018127.6(ELAC2):c.2342G>A (p.Arg781His)	GGAGGAGCRCAGGGAGAAGCGGG	Prostate cancer, hereditary, 2
137854450	ELANE	NM_001972.2(ELANE):c.377C>T (p.Ser126Leu)	CCGCCACAGCTCAACGGGTYGGC, CCACAGCTCAACGGGTYGGCCAC	Severe congenital neutropenia autosomal dominant
727503035	ELN	NM_000501.3(ELN):c.1918+1G>A	GTTTGRTGAGCACTGGGTGGAGG, CCAGTTTGRTGAGCACTGGGTGG	Supravalvar aortic stenosis
137854452	ELN	NM_000501.3(ELN):c.1324C>T (p.Gln442Ter)	CCTTGTAGCCGAAGCTYAGGCAG	Supravalvar aortic stenosis
515726212	EPB42	NM_000119.2(EPB42):c.949C>T (p.Arg317Cys)	CCACGCRGGCAGGGATTCCCAGG	Spherocytosis type 5
267606785	EPCAM	NM_002354.2(EPCAM):c.197G>A (p.Cys66Tyr)	TGCCAAATRTTTGGTGATGAAGG	Diarrhea 5, with tufting enteropathy, congenital
116506614	EPHA2	NM_004431.3(EPHA2):c.2162G>A (p.Arg721Gln)	CCAGCTGCGATGCCCYGCAGCAT	
28933368	ERBB2	NM_001005862.2(ERBB2):c.2650G>A (p.Glu884Lys)	GTGGRAGCTGATGACTTTTGGGG, TGTGGRAGCTGATGACTTTTGGG,	Glioma susceptibility 1

			GTGTGGRAGCTGATGACTTTTGG	
121913023	ERCC2	NM_000400.3(ERCC2):c.2041G>A (p.Asp681Asn)	GCCRACAAGGTGCAGCTTCAGGG, TGCCRACAAGGTGCAGCTTCAGG	Cerebrooculofacioskeletal syndrome 2
41556519	ERCC2	NM_000400.3(ERCC2):c.2047C>T (p.Arg683Trp)	CCTGCGCTTCTGCCCACAGYGGT	Xeroderma pigmentosum, group D
121913017	ERCC2	NM_000400.3(ERCC2):c.2176C>T (p.Gln726Ter)	CCTGCGGCAGATGGCAYAGCCCT	Xeroderma pigmentosum, group D
121913021	ERCC2	NM_000400.3(ERCC2):c.1972C>T (p.Arg658Cys)	CCTTCGATGCCATGYGCCACGCG	Photosensitive trichothiodystrophy
121913024	ERCC2	NM_000400.3(ERCC2):c.1846C>T (p.Arg616Trp)	CCAGTGCACCACTACGGGYGGGC	Cerebro-oculo-facio-skeletal syndrome, Xeroderma pigmentosum, group D, Cerebrooculofacioskeletal syndrome 2
121913026	ERCC2	NM_000400.3(ERCC2):c.2164C>T (p.Arg722Trp)	CCAAGTACTTCCTGYGGCAGATG	Photosensitive trichothiodystrophy
121913047	ERCC3	NM_000122.1(ERCC3):c.1273C>T (p.Arg425Ter)	CCTGGGAGGCCGAGYGAGTCATG	Xeroderma pigmentosum, complementation group b
121917904	ERCC6	NM_000124.3(ERCC6):c.2047C>T (p.Arg683Ter)	CCGATGCAAAATAACCTCYGAGA	Cerebro-oculo-facio-skeletal syndrome
121434325	ERCC8	NM_000082.3(ERCC8):c.479C>T (p.Ala160Val)	CCAAGCACTGTTTGGTAGYAGGT	Cockayne syndrome type A, not provided
587777006	ERF	NM_006494.3(ERF):c.547C>T (p.Arg183Ter)	TGAGCCTCRGCCCAGGCGGCGGG	Craniosynostosis 4
587777010	ERF	NM_006494.3(ERF):c.1270C>T (p.Gln424Ter)	TCTRTGGTGGCGGGGGCGGTGGG, ATCTRTGGTGGCGGGGGCGGTGG, TTGATCTRTGGTGGCGGGGGCGG	Craniosynostosis 4
56025238	ERMAP	NM_001017922.1(ERMAP):c.169G>A (p.Gly57Arg)	GCCCRGGACGGTACCCAAGGAGG, CTGGCCCRGGACGGTACCCAAGG	
80359865	ESCO2	NM_001017420.2(ESCO2):c.1354-18G>A	TCTTRGTTTTTAAAATCATTAGG	Roberts-SC phocomelia syndrome
80359850	ESCO2	NM_001017420.2(ESCO2):c.604C>T (p.Gln202Ter)	CCAAAAAATAAAACCAYAAGTTA	Roberts-SC phocomelia syndrome
121908136	ESPN	NM_031475.2(ESPN):c.2321G>A (p.Arg774Gln)	GGCRGAAGGTGGGTGGGGCGGGG, AGGCRGAAGGTGGGTGGGGCGGG, GAGGCRGAAGGTGGGTGGGGCG	Deafness, without vestibular involvement, autosomal dominant

			G, GCAGAGGCRGAAGGTGGGTGGGG	
104893956	ESR1	NM_001122742.1(ESR1):c.469C>T (p.Arg157Ter)	CCAAATTCAGATAATYGACGCCA	Estrogen resistance
119458971	ETFA	NM_000126.3(ETFA):c.346G>A (p.Gly116Arg)	CTGCCTTCRGAAAGGTGAGAAGG	Glutaric acidemia IIA
104894677	ETFB	NM_001985.2(ETFB):c.491G>A (p.Arg164Gln)	GTGGAGCRGGAGATCGATGGGGG, AGTGGAGCRGGAGATCGATGGGG	Glutaric acidemia IIB
796051960	ETFDH	NM_004453.3(ETFDH):c.1809G>A (p.Trp603Ter)	AACTGRGTGGTACCTGAAGGTGG, ATTAACTGRGTGGTACCTGAAGG	not provided
724159946	ETV6	NM_001987.4(ETV6):c.1106G>A (p.Arg369Gln)	TTCCRGATAGTGGATCCCAACGG	Hematologic neoplasm, Thrombocytopenia, Thrombocytopenia 5
724159947	ETV6	NM_001987.4(ETV6):c.641C>T (p.Pro214Leu)	CCGCCGCCTCTCCCYGGCTGAGA	Hematologic neoplasm, Thrombocytopenia, Thrombocytopenia 5
121909199	EYA1	NM_000503.5(EYA1):c.1276G>A (p.Gly426Ser)	GTGTACGGRGCGGTGTGGACTGG	
527236064	EYS	NM_001142800.1(EYS):c.7793G>A (p.Gly2598Asp)	CCTGAGGRCCACCCAAATGCTGG	Retinitis pigmentosa
794727631	EYS	NM_001142800.1(EYS):c.490C>T (p.Arg164Ter)	CCTTGTCCACTGGGACTTYGACT	Retinitis pigmentosa 25
587783625	EZH2	NM_004456.4(EZH2):c.1876G>A (p.Val626Met)	CATCTGACRTGGCAGGCTGGGGG	Weaver syndrome
61753266	F10	NM_000504.3(F10):c.424G>A (p.Glu142Lys)	CCACRAGGAACAGAACTCTGTGG	Factor X deficiency
121913071	F13A1	NM_000129.3(F13A1):c.782G>A (p.Arg261His)	CAGCCRTGTGGGGTCTGCAATGG	Factor xiii, a subunit, deficiency of
121913065	F13A1	NM_000129.3(F13A1):c.514C>T (p.Arg172Ter)	CCCTATGGCGTACTTYGAACCAG, CCTATGGCGTACTTYGAACCAGT	Factor xiii, a subunit, deficiency of
267606787	F13A1	NM_000129.3(F13A1):c.2110C>T (p.Arg704Trp)	CCCTGGGTCTCTGGGCATYGGAA, CCTGGGTCTCTGGGCATYGGAAG	Factor xiii, a subunit, deficiency of
267606789	F13A1	NM_000129.3(F13A1):c.1984C>T (p.Arg662Ter)	CCTTTAAAAGAAACCCTGYGAAA	Factor xiii, a subunit, deficiency of
121918482	F2	NM_000506.3(F2):c.1292G>A (p.Arg431His)	TCCCRCACCAGGTACAGAACTGG	
121918483	F2	NM_000506.3(F2):c.1027G	GTTCRAGAAGAAGTCGCTGGAGG	Hereditary factor II deficiency disease

		>A (p.Glu343Lys)	, TCTGTTCRAGAAGAAGTCGCTGG	
121918484	F2	NM_000506.3(F2):c.1054G >A (p.Glu352Lys)	CAAAACCRAAAGAGAGCTCCTGG	Hereditary factor II deficiency disease
386834228	F5	NM_000130.4(F5):c.5668G >A (p.Glu1890Lys)	TTGGARAAAACCAGAGAGCAGG G, GTTGGARAAAACCAGAGAGCAG G	not provided
36209567	F7	NM_000131.4(F7):c.1061C >T (p.Ala354Val)	CCGTGGCGCCACGGYCCTGGAGC	Factor VII deficiency
137852373	F8	NM_000132.3(F8):c.5167G >A (p.Glu1723Lys)	GTGRAGAGGCTCTGGGATTATGG	Hereditary factor VIII deficiency disease
137852466	F8	NM_000132.3(F8):c.6545G >A (p.Arg2182His)	ACTCTTCRCATGGAGTTGATGGG, CACTCTTCRCATGGAGTTGATGG	Hereditary factor VIII deficiency disease
28933681	F8	NM_000132.3(F8):c.5710G >A (p.Glu1904Lys)	TCTTTGATRAGACCAAAAGCTGG	Hereditary factor VIII deficiency disease
137852357	F8	NM_000132.3(F8):c.6496C >T (p.Arg2166Ter)	CCCTCCAATTATTGCTYGATACA, CCTCCAATTATTGCTYGATACAT	Hereditary factor VIII deficiency disease
137852364	F8	NM_000132.3(F8):c.1171C >T (p.Arg391Cys)	CCTTCCTTTATCCAAATTYGCTC, CCTTTATCCAAATTYGCTCAGTT	Hereditary factor VIII deficiency disease
137852368	F8	NM_000132.3(F8):c.1063C >T (p.Arg355Ter)	CCAGAGGAACCCCAACTAYGAAT	Hereditary factor VIII deficiency disease
137852393	F8	NM_000132.3(F8):c.493C> T (p.Pro165Ser)	CCTGAAAGAGAATGGTYCAATGG	Hereditary factor VIII deficiency disease
137852401	F8	NM_000132.3(F8):c.881C> T (p.Thr294Ile)	CCTCGAAGGTCACAYATTTCTTG	Hereditary factor VIII deficiency disease
137852416	F8	NM_000132.3(F8):c.1636C >T (p.Arg546Trp)	CCAACTAAATCAGATCCTYGGTG	Hereditary factor VIII deficiency disease
137852428	F8	NM_000132.3(F8):c.1834C >T (p.Arg612Cys)	CCTCACAGAGAATATACAAYGCT	Hereditary factor VIII deficiency disease
137852435	F8	NM_000132.3(F8):c.2149C >T (p.Arg717Trp)	CCACAACTCAGACTTTYGGAACA	Hereditary factor VIII deficiency disease
137852445	F8	NM_000132.3(F8):c.5422C >T (p.Leu1808Phe)	CCTTCTATTCTAGCYTTATTTCT	Hereditary factor VIII deficiency disease
137852453	F8	NM_000132.3(F8):c.6046C >T (p.Arg2016Trp)	CCAAAGCTGGAATTTGGYGGGTG	Hereditary factor VIII deficiency disease
137852456	F8	NM_000132.3(F8):c.6263C >T (p.Ser2088Phe)	CCAAGGAGCCCTTTTYTTGGATC	Hereditary factor VIII deficiency disease
137852463	F8	NM_000132.3(F8):c.6518C	CCGTTTGCACCCAAYTCATTATA	Hereditary factor VIII deficiency disease

		>T (p.Thr2173Ile)		
137852464	F8	NM_000132.3(F8):c.6532C>T (p.Arg2178Cys)	CCCAACTCATTATAGCATTYGCA, CCAACTCATTATAGCATTYGCAG	Hereditary factor VIII deficiency disease
137852473	F8	NM_000132.3(F8):c.6967C>T (p.Arg2323Cys)	CCCACCGTTACTGACTYGCTACC, CCACCGTTACTGACTYGCTACCT	Hereditary factor VIII deficiency disease
137852257	F9	NM_000133.3(F9):c.1069G>A (p.Gly357Arg)	TGGRGAAGAGTCTTCCACAAAGG	Hereditary factor IX deficiency disease
137852267	F9	NM_000133.3(F9):c.1324G>A (p.Gly442Arg)	CAAATATRGAATATATACCAAGG	Hereditary factor IX deficiency disease
137852275	F9	NM_000133.3(F9):c.1070G>A (p.Gly357Glu)	GGGRAAGAGTCTTCCACAAAGGG, TGGGRAAGAGTCTTCCACAAAGG	Hereditary factor IX deficiency disease
137852272	F9	NM_000133.3(F9):c.484C>T (p.Arg162Ter)	CCTGTACTGAGGGATATYGACTT	Hereditary factor IX deficiency disease
387907040	FA2H	NM_024306.4(FA2H):c.460C>T (p.Arg154Cys)	CCGGTGACCAGGCCCATCYGCCT	Spastic paraplegia 35
80338901	FAH	NM_000137.2(FAH):c.1062+5G>A	TGARTATCTGGCTGCACTGAGGG, GTGARTATCTGGCTGCACTGAGG	Tyrosinemia type I, not provided
587777011	FAM111A	NM_001142519.1(FAM111A):c.1706G>A (p.Arg569His)	ACTCRTAGTATCATTGAGTTTGG	Kenny-Caffey syndrome type 2
587777238	FAM111B	NM_198947.3(FAM111B):c.1883G>A (p.Ser628Asn)	GAARTTTCCTATCAGAGGTTTGG, CCAAAGAARTTTCCTATCAGAGG	Poikiloderma, hereditary fibrosing, with tendon contractures, myopathy, and pulmonary fibrosis
137852737	FAM134B	NM_001034850.2(FAM134B):c.433C>T (p.Gln145Ter)	CCCTGTTGCAGGTGCAYAGTTGT, CCTGTTGCAGGTGCAYAGTTGTG	Hereditary sensory and autonomic neuropathy type IIB, Hereditary sensory and autonomic neuropathy type IIA
796051850	FAM20C	NM_020223.3(FAM20C):c.1645C>T (p.Arg549Trp)	CCTGGAGGCCCTGGACCGGYGGC	Raine syndrome
137854435	FAM83H	NM_198488.3(FAM83H):c.973C>T (p.Arg325Ter)	CCCCTTCTCCTTCCCTAAAYGAG, CCCTTCTCCTTCCCTAAAYGAGC, CCTTCTCCTTCCCTAAAYGAGCG	Amelogenesis imperfecta, hypocalcification type
137854440	FAM83H	NM_198488.3(FAM83H):c.2029C>T (p.Gln677Ter)	CCTGAACCCCCTGGTCYAGCGCA	Amelogenesis imperfecta, hypocalcification type
387907056	FAM83H	NM_198488.3(FAM83H):c.1366C>T (p.Gln456Ter)	CCGTGACCAGCTCTACCAGYAGC	Amelogenesis imperfecta, hypocalcification type
730881731	FANCC	NM_000136.2(FANCC):c.319C>T (p.Gln107Ter)	CCACAGAATTCTGGAYAATCAAA	Hereditary cancer-predisposing syndrome
121434506	FANCE	NM_021922.2(FANCE):c.421C>T (p.Arg141Ter)	CCCTGGGGGAATTGCTGYGAAGG, CCTGGGGGAATTGCTGYGAAGGG	Fanconi anemia, complementation group E

121918163	FANCI	NM_001113378.1(FANCI):c.3854G>A (p.Arg1285Gln)	TCACRAGACTTCAAGATCAAAGG	Fanconi anemia, complementation group I
121913077	FAS	NM_000043.4(FAS):c.817C>T (p.Gln273Ter)	CCAAGACACAGCAGAAYAGAAAG	Autoimmune lymphoproliferative syndrome, type 1a
398122955	FAT4	NM_024582.4(FAT4):c.7123G>A (p.Glu2375Lys)	ATTCCTRAGGATGCACCAACTGG	Van Maldergem syndrome 2, Hennekam lymphangiectasia-lymphedema syndrome 2
80338765	FBLN5	NM_006329.3(FBLN5):c.604G>A (p.Gly202Arg)	GAGGATRGAAGGTCTTGCCAAGG	Autosomal recessive cutis laxa type IA
28939073	FBLN5	NM_006329.3(FBLN5):c.1051C>T (p.Arg351Trp)	CCCTTTACCATCTTGTACYGGGA, CCTTTACCATCTTGTACYGGGAC	Age-related macular degeneration 3
193922236	FBN1	NM_000138.4(FBN1):c.7806G>A (p.Trp2602Ter)	TACCAGTGRAACCAGTGTGTTGG	Thoracic aortic aneurysms and aortic dissections, Marfan syndrome
794728166	FBN1	NM_000138.4(FBN1):c.1421G>A (p.Cys474Tyr)	ACCGGTRTGAGTGCAACAAAGGG, TACCGGTRTGAGTGCAACAAAGG	Thoracic aortic aneurysms and aortic dissections
794728170	FBN1	NM_000138.4(FBN1):c.1583G>A (p.Cys528Tyr)	AGAATRCCGAGGTATGGTCCTGG	Thoracic aortic aneurysms and aortic dissections
794728237	FBN1	NM_000138.4(FBN1):c.5699G>A (p.Cys1900Tyr)	CCTRTGGGAATGGAACTTGCCGG	Thoracic aortic aneurysms and aortic dissections
794728240	FBN1	NM_000138.4(FBN1):c.5801G>A (p.Cys1934Tyr)	ATGAATRTGCAAGTGGAAATGGG, GATGAATRTGCAAGTGGAAATGG	Thoracic aortic aneurysms and aortic dissections
794728257	FBN1	NM_000138.4(FBN1):c.6871G>A (p.Asp2291Asn)	TGTARGTAAGAGGATCCCTGTGG	Thoracic aortic aneurysms and aortic dissections
794728266	FBN1	NM_000138.4(FBN1):c.7205-1G>A	TACARATATCGATGAATGCAAGG	Thoracic aortic aneurysms and aortic dissections
397515757	FBN1	NM_000138.4(FBN1):c.1468+5G>A	GTACRTGATCCATCCTAGGTTGG, ATTGGTACRTGATCCATCCTAGG	Thoracic aortic aneurysms and aortic dissections, Marfan syndrome
397515804	FBN1	NM_000138.4(FBN1):c.4259G>A (p.Cys1420Tyr)	CAGTRCCTCAATGCACCAGGAGG, GGCCAGTRCCTCAATGCACCAGG	Thoracic aortic aneurysms and aortic dissections, Marfan syndrome
397515859	FBN1	NM_000138.4(FBN1):c.7955G>A (p.Cys2652Tyr)	CAATGAATRTGGCTCTGCGCAGG	Thoracic aortic aneurysms and aortic dissections, Marfan syndrome
775417975	FBN1	NM_000138.4(FBN1):c.3513C>A (p.Cys1171Ter)	GGTTCACRCAACGGCCATTGGGG, AGGTTCACRCAACGGCCATTGGG	Thoracic aortic aneurysms and aortic dissections
794728335	FBN1	NM_000138.4(FBN1):c.6425G>A (p.Cys2142Tyr)	GGACAGTRCATCAATACAGATGG	Thoracic aortic aneurysms and aortic dissections
548296552	FBN1	NM_000138.4(FBN1):c.292	ATGCRGTGGCGGCCAGCAATAGG	Thoracic aortic aneurysms and aortic dissections

		6C>T (p.Arg976Cys)		
137854475	FBN1	NM_000138.4(FBN1):c.3509G>A (p.Arg1170His)	GCCRTTGCGTGAACCTCATAGGG, GGCCRTTGCGTGAACCTCATAGG	Thoracic aortic aneurysms and aortic dissections, Marfan syndrome, Marfanoid habitus, not specified, not provided
137854482	FBN1	NM_000138.4(FBN1):c.3386G>A (p.Cys1129Tyr)	GGTGTTTRCCATAACACAGAGGG , TGGTGTTTRCCATAACACAGAGG	Marfan syndrome
137854483	FBN1	NM_000138.4(FBN1):c.3662G>A (p.Cys1221Tyr)	TATGAATRTAGCTGTCAGCCGGG, CTATGAATRTAGCTGTCAGCCGG	Marfan syndrome
137854484	FBN1	NM_000138.4(FBN1):c.3257G>A (p.Cys1086Tyr)	CCAGTRTGTGAACACCCCTGGGG, GCCAGTRTGTGAACACCCCTGGG, GGCCAGTRTGTGAACACCCCTGG	
369294972	FBN1	NM_000138.4(FBN1):c.7660C>T (p.Arg2554Trp)	GAATCCCCRCTGGCATTCACAGG	Thoracic aortic aneurysms and aortic dissections, Marfan syndrome, incomplete
113871094	FBN1	NM_000138.4(FBN1):c.4786C>T (p.Arg1596Ter)	CCTGGAGGGGAAGGTTTCYGACC	Thoracic aortic aneurysms and aortic dissections, Marfan syndrome
794728262	FBN1	NM_000138.4(FBN1):c.7003C>T (p.Arg2335Trp)	CCCTTCCAGACAATYGGGAAGGG	Thoracic aortic aneurysms and aortic dissections
730880099	FBN1	NM_000138.4(FBN1):c.1633C>T (p.Arg545Cys)	CCGGATCTGCAATAATGGAYGCT	Marfan syndrome
794728195	FBN1	NM_000138.4(FBN1):c.2645C>T (p.Ala882Val)	CCTCCCTCGGTGCTGYGTGGGGA	Thoracic aortic aneurysms and aortic dissections
794728196	FBN1	NM_000138.4(FBN1):c.2671C>T (p.Gln891Ter)	CCCGTGCACCCTATGCYAAGTTG, CCGTGCACCCTATGCYAAGTTGG	Thoracic aortic aneurysms and aortic dissections
794728231	FBN1	NM_000138.4(FBN1):c.4888C>T (p.Gln1630Ter)	CCTTTGGGAGTTTCYAGTGCCGC	Thoracic aortic aneurysms and aortic dissections
397514558	FBN1	NM_000138.4(FBN1):c.2920C>T (p.Arg974Cys)	CCCTGCCTATTGCTGGCYGCCAC, CCTGCCTATTGCTGGCYGCCACC	Marfan syndrome, Ectopia lentis, isolated, autosomal dominant
794728283	FBN1	NM_000138.4(FBN1):c.8038C>T (p.Arg2680Cys)	CCACCTGGTTACTTCYGCATAGG	Thoracic aortic aneurysms and aortic dissections
140630	FBN1	NM_000138.4(FBN1):c.4930C>T (p.Arg1644Ter)	CCTGAATGAAGATACAYGAGTGT	Thoracic aortic aneurysms and aortic dissections
113001196	FBN1	NM_000138.4(FBN1):c.6658C>T (p.Arg2220Ter)	CCTCTGCTCTGTGCCTTCYGATG	Thoracic aortic aneurysms and aortic dissections, Marfan syndrome
112645512	FBN1	NM_000138.4(FBN1):c.1285C>T (p.Arg429Ter)	CCTCAAATTCCGGTCCCTYGACC	Marfan syndrome
25403	FBN1	NM_000138.4(FBN1):c.184C>T (p.Arg62Cys)	CCCAATGTCTGTGGATCAYGTTA, CCAATGTCTGTGGATCAYGTTAT	Marfan syndrome
137852826	FBN2	NM_001999.3(FBN2):c.1171G>A (p.Glu391Lys)	TGTRAGCCTGGCCGCTGCTGGGG, CTGTRAGCCTGGCCGCTGCTGGG,	Congenital contractural arachnodactyly

			GCTGTRAGCCTGGCCGCTGCTGG	
121918188	FBP1	NM_001127628.1(FBP1):c.490G>A (p.Gly164Ser)	GCAGCCRGCTACGCACTGTATGG	Fructose-biphosphatase deficiency
398123061	FBXL4	NM_012160.4(FBXL4):c.1444C>T (p.Arg482Trp)	CCAAGTGTAAAAAACTCYGGACC	Mitochondrial encephalomyopathy, Mitochondrial DNA depletion syndrome 13 (encephalomyopathic type), Global developmental delay
121918305	FBXO7	NM_012179.3(FBXO7):c.65C>T (p.Thr22Met)	CCCGAGACGGAGCCGAYGCTGGG, CCGAGACGGAGCCGAYGCTGGGG	Parkinson disease 15
267606804	FECH	NM_001012515.2(FECH):c.1243C>T (p.Pro415Ser)	CCGCTCTGTGTCAATYCTGTCTG	Erythropoietic protoporphyria
121918292	FERMT1	NM_017671.4(FERMT1):c.787C>T (p.Gln263Ter)	CCTTATGGAACAAGGCATCYAAG	Kindler syndrome
121918296	FERMT3	NM_178443.2(FERMT3):c.48G>A (p.Trp16Ter)	ATGRGAGCTGCGGGTGTTTGTGG	LEUKOCYTE ADHESION DEFICIENCY, TYPE III
121918298	FERMT3	NM_178443.2(FERMT3):c.687G>A (p.Trp229Ter)	CCAGGTGRCTGGACTCGTCGCGG	LEUKOCYTE ADHESION DEFICIENCY, TYPE III
121918295	FERMT3	NM_178443.2(FERMT3):c.1537C>T (p.Arg513Ter)	CCCCCCGTTTCCAGYGAAAGTTC	LEUKOCYTE ADHESION DEFICIENCY, TYPE III
121918297	FERMT3	NM_178443.2(FERMT3):c.1729C>T (p.Arg577Ter)	CCTGGGCATCGCCAACAACYGAC	LEUKOCYTE ADHESION DEFICIENCY, TYPE III
121909607	FGA	NM_000508.3(FGA):c.104G>A (p.Arg35His)	CGTGCRTGGCCCAAGGGTTGTGG	Dysfibrinogenemia
606231223	FGB	NM_005141.4(FGB):c.958+13C>T	CCAGGTAACGAACAGGYATGCAA	Afibrinogenemia, congenital
28935498	FGD1	NM_004463.2(FGD1):c.935C>T (p.Pro312Leu)	CCCAGCCACAGCCTCTGCCYTGG, CCAGCCACAGCCTCTGCCYTGGG, CCACAGCCTCTGCCYTGGGCCCC	Syndromic X-linked mental retardation 16
387906718	FGD1	NM_004463.2(FGD1):c.1966C>T (p.Arg656Ter)	CCAACCTCAATCTGCCTYGAACC	Aarskog syndrome
118203974	FGD4	NM_139241.3(FGD4):c.823C>T (p.Arg275Ter)	CCAGAGCTGGAGAAAYGAATGCA	Charcot-Marie-Tooth disease, type 4H
121917704	FGF3	NM_005247.2(FGF3):c.310C>T (p.Arg104Ter)	CCATGAACAAGAGGGGAYGACTC	Deafness with labyrinthine aplasia microtia and microdontia (LAMM)
137852660	FGF8	NM_033163.3(FGF8):c.77C>T (p.Pro26Leu)	CCTCTAGGAAGGCCYGGGCAGGG	Kallmann syndrome 6
121918322	FGF9	NM_002010.2(FGF9):c.296G>A (p.Ser99Asn)	TATCARTATAGCAGTGGGCCTGG	Multiple synostoses syndrome 3

515726225	FGFR1	NM_023110.2(FGFR1):c.2084C>T (p.Thr695Ile)	CCAGARTGAAGATCTCCCACAGG	Kallmann syndrome 2
121909636	FGFR1	NM_023110.2(FGFR1):c.2038C>T (p.Gln680Ter)	CCGGATCTACACCCACYAGAGTG	Kallmann syndrome 2, Delayed puberty
515726224	FGFR1	NM_023110.2(FGFR1):c.1460G>A (p.Gly487Asp)	CCACCTGCCCAAAGCAGYCCTCT, CCTGCCCAAAGCAGYCCTCTCCC	Kallmann syndrome 2
121918491	FGFR2	NM_000141.4(FGFR2):c.1032G>A (p.Ala344=)	GCTTGGCRGGTAATTCTATTGGG, TGCTTGGCRGGTAATTCTATTGG	Crouzon syndrome, Craniosynostosis
121918509	FGFR2	NM_000141.4(FGFR2):c.1882G>A (p.Ala628Thr)	TTTARCAGCCAGAAATGTTTTGG	
121913112	FGFR3	NM_000142.4(FGFR3):c.1537G>A (p.Asp513Asn)	CACAGACRATGCCACTGACAAGG	
351855	FGFR4	NM_213647.2(FGFR4):c.1162G>A (p.Gly388Arg)	CCTGCCCTCGATACAGCCYGGCC, CCCTCGATACAGCCYGGCCAGCA	
104894689	FKRP	NM_024301.4(FKRP):c.764G>A (p.Trp255Ter)	CGCTRGAAGGCTGAGCGCGAGGG, GCGCTRGAAGGCTGAGCGCGAGG	Limb-girdle muscular dystrophy-dystroglycanopathy, type C5
104894681	FKRP	NM_024301.4(FKRP):c.1343C>T (p.Pro448Leu)	CCCGAGCACTTCCTGCAGCYGCT, CCGAGCACTTCCTGCAGCYGCTG	Congenital muscular dystrophy-dystroglycanopathy without mental retardation, type B5
104894690	FKRP	NM_024301.4(FKRP):c.400C>T (p.Arg134Trp)	CCTAGTACCTGATGGGGCGYGGG	Limb-girdle muscular dystrophy-dystroglycanopathy, type C5
587782069	FLCN	NM_144997.5(FLCN):c.499C>T (p.Gln167Ter)	CCTGGCCAGGGGCTTCYAGCGCT	Hereditary cancer-predisposing syndrome
398124523	FLCN	NM_144997.5(FLCN):c.1060C>T (p.Gln354Ter)	CCCTCCGGCACATGAGGYAGGTA, CCTCCGGCACATGAGGYAGGTAG	not provided
398124532	FLCN	NM_144997.5(FLCN):c.1597C>T (p.Gln533Ter)	CCCAAAGAGGACACAYAGAAGCT, CCAAAGAGGACACAYAGAAGCTG	not provided
387907371	FLNA	NM_001110556.1(FLNA):c.5217G>A (p.Thr1739=)	TGACRGTGAGGAGGGGTGGGGG, GTGACRGTGAGGAGGGGTGGGGG, AGTGACRGTGAGGAGGGGTGGGG, AAGTGACRGTGAGGAGGGGTGGG, CAAGTGACRGTGAGGAGGGGTG	Terminal osseous dysplasia

			G	
28935473	FLNA	NM_001110556.1(FLNA):c.3596C>T (p.Ser1199Leu)	CCATTGAGATCTGCTYGGAGGCG	Melnick-Needles syndrome
80338841	FLNA	NM_001110556.1(FLNA):c.1923C>T (p.Gly641=)	CCGCAGGAGGCTGGYGAGTATGC	X-linked periventricular heterotopia, Cardiac valvular dysplasia, X-linked
398123614	FLNA	NM_001110556.1(FLNA):c.2761C>T (p.Arg921Ter)	CCAAGGGGGATGCAGTGYGAGAT	X-linked periventricular heterotopia, Oto-palato-digital syndrome, type I, not provided
137853310	FLNA	NM_001110556.1(FLNA):c.544C>T (p.Gln182Ter)	CCAGAACAAGCTGCCGYAGCTGC	X-linked periventricular heterotopia
137853312	FLNA	NM_001110556.1(FLNA):c.3557C>T (p.Ser1186Leu)	CCAAGTGGACTGCTYGAGCGCGG	Frontometaphyseal dysplasia
137853317	FLNA	NM_001110556.1(FLNA):c.586C>T (p.Arg196Trp)	CCGGGACTGGCAGAGCGGCYGGG	Oto-palato-digital syndrome, type I, Oto-palato-digital syndrome, type II, not provided
80356510	FLNB	NM_001457.3(FLNB):c.1088G>A (p.Gly363Glu)	GTCCAGRGTTGGAAGCTGTAGGG, GGTCCAGRGTTGGAAGCTGTAGG	Larsen syndrome, dominant type
80356513	FLNB	NM_001457.3(FLNB):c.4756G>A (p.Gly1586Arg)	AAGACTRGGCGCTATATGATTGG	Larsen syndrome, dominant type, Larsen syndrome
80356517	FLNB	NM_001457.3(FLNB):c.1945C>T (p.Arg649Ter)	CCTTTGCTTCAGGTTYGAGCATA	Spondylocarpotarsal synostosis syndrome
80356519	FLNB	NM_001457.3(FLNB):c.2452C>T (p.Arg818Ter)	CCTCCTGCTGCTGGGYGATACAC	Spondylocarpotarsal synostosis syndrome
121909654	FLT4	NM_182925.4(FLT4):c.2632G>A (p.Val878Met)	GTGGCCRTGAAAATGCTGAAAGG	Hereditary lymphedema type I
121909656	FLT4	NM_182925.4(FLT4):c.3316G>A (p.Glu1106Lys)	CTCTGGRAGATCTTCTCTCTGGG, TCTCTGGRAGATCTTCTCTCTGG	Hereditary lymphedema type I
121909657	FLT4	NM_182925.4(FLT4):c.2563G>A (p.Ala855Thr)	CGGCRCCTTCGGGAAGGTGGTGG, CTACGGCRCCTTCGGGAAGGTGG	Hereditary lymphedema type I
34255532	FLT4	NM_182925.4(FLT4):c.2860C>T (p.Pro954Ser)	CCGCAGGAGAAGTCTYCCGAGCA	Hemangioma, capillary infantile
267606819	FLVCR1	NM_014053.3(FLVCR1):c.721G>A (p.Ala241Thr)	CACCRCCGTGCTGGGCAATCAGG	Posterior column ataxia with retinitis pigmentosa
72549320	FMO3	NM_001002294.2(FMO3):c.94G>A (p.Glu32Lys)	TTTRAGAAGAGCAATGACATTGG	Trimethylaminuria
2266782	FMO3	NM_006894.5(FMO3):c.472G>A (p.Glu158Lys)	AAAARAGTCCTTTCCAGGTAAGG	Trimethylaminuria
72549326	FMO3	NM_006894.5(FMO3):c.458C>T (p.Pro153Leu)	CCGGACATCATGTGTATCYCAAC	Trimethylaminuria

79691946	FOXC1	NM_001453.2(FOXC1):c.889C>T (p.Pro297Ser)	CCGCCGCCGCCCGCGYCCTCCGC	Iridogoniodysgenesis type1, not specified, not provided
104893952	FOXC1	NM_001453.2(FOXC1):c.67C>T (p.Gln23Ter)	CCCTACCTCGGCGGCGAGYAGAG, CCTACCTCGGCGGCGAGYAGAGC	Axenfeld-Rieger syndrome type 3
104893957	FOXC1	NM_001453.2(FOXC1):c.392C>T (p.Ser131Leu)	CCGCCACAACCTCTYGCTCAACG	Axenfeld-Rieger syndrome type 3
786205000	FOXG1	NM_005249.4(FOXG1):c.136C>T (p.Gln46Ter)	CCACAACAGCCACCACCCCYAGC	not provided
786205006	FOXG1	NM_005249.4(FOXG1):c.610C>T (p.Leu204Phe)	CCCCGAGAAGCGGCTCACGYTCA, CCCGAGAAGCGGCTCACGYTCAA, CCGAGAAGCGGCTCACGYTCAAC	not provided
796052467	FOXG1	NM_005249.4(FOXG1):c.701C>T (p.Ser234Phe)	CCATCCGCCACAATCTGTYCCTC, CCGCCACAATCTGTYCCTCAACA	not provided
796052458	FOXG1	NM_005249.4(FOXG1):c.217C>T (p.Gln73Ter)	CCGCCGCCGCAGCAGCAGYAGCC, CCGCCGCAGCAGCAGYAGCCGCC	not provided
387906920	FOXL2	NM_023067.3(FOXL2):c.205G>A (p.Glu69Lys)	TCCGCRAGAGCGCGGAGAAGAGG	
104893739	FOXL2	NM_023067.3(FOXL2):c.586C>T (p.Gln196Ter)	CCCCCCAAGTACCTGYAGTCTGG, CCCCCAAGTACCTGYAGTCTGGC	Blepharophimosis, ptosis, and epicanthus inversus
104893741	FOXL2	NM_023067.3(FOXL2):c.655C>T (p.Gln219Ter)	CCCTATGCCTCCTGCYAGATGGC, CCTATGCCTCCTGCYAGATGGCG	Blepharophimosis, ptosis, and epicanthus inversus
122467174	FOXP3	NM_014009.3(FOXP3):c.3G>A (p.Met1Ile)	CCGATRCCCAACCCCAGGCCTGG	Insulin-dependent diabetes mellitus secretory diarrhea syndrome
120074156	FRAS1	NM_025074.6(FRAS1):c.8602C>T (p.Gln2868Ter)	CCTGGTGTCATTGAAYAGGTGCG	Cryptophthalmos syndrome
137852209	FRMD7	NM_194277.2(FRMD7):c.252G>A (p.Val84=)	AGTRGACCCTGGACATCTGCGGG, CAGTRGACCCTGGACATCTGCGG	Infantile nystagmus, X-linked
137852208	FRMD7	NM_194277.2(FRMD7):c.1003C>T (p.Arg335Ter)	CCCATCTCAGTACCATGAAYGAC, CCATCTCAGTACCATGAAYGACA	Infantile nystagmus, X-linked
121909660	FSHR	NM_000145.3(FSHR):c.1717C>T (p.Arg573Cys)	CCAGGATCGCCAAGYGCATGGCC	Ovarian dysgenesis 1
28941768	FTCD	NM_006657.2(FTCD):c.403C>T (p.Arg135Cys)	CCAGGATGGACAGTYGCCGGACC	Glutamate formiminotransferase deficiency
104894685	FTL	NM_000146.3(FTL):c.286G	AGACRCCATGAAAGCTGCCATGG	Neuroferritinopathy

		>A (p.Ala96Thr)		
397514540	FTL	NM_000146.3(FTL):c.89C>T (p.Thr30Ile)	CCTGCAGGCCTCCTACAYCTACC	Hyperferritinemia cataract syndrome
121909669	FUS	NM_004960.3(FUS):c.1553G>A (p.Arg518Lys)	ACARACAGGATCGCAGGGAGAGG, TGAGCACARACAGGATCGCAGGG	Amyotrophic lateral sclerosis type 6
267606831	FUS	NM_004960.3(FUS):c.1520G>A (p.Gly507Asp)	TGGCTTTGRCCCTGGCAAGATGG	Amyotrophic lateral sclerosis type 6
387906628	FUS	NM_004960.3(FUS):c.616G>A (p.Gly206Ser)	CAGCRGTGGCTATGGACAGCAGG	Amyotrophic lateral sclerosis type 6
104894569	G6PC	NM_000151.3(G6PC):c.551G>A (p.Gly184Glu)	GCTGRAGTCCTGTCAGGTATGGG, TGCTGRAGTCCTGTCAGGTATGG	Glycogen storage disease type 1A
1801176	G6PC	NM_000151.3(G6PC):c.248G>A (p.Arg83His)	TGGACAGCRTCCATACTGGTGGG	Glucose-6-phosphate transport defect
137852316	G6PD	NM_000402.4(G6PD):c.1268G>A (p.Arg423His)	GATCCRCGTGCAGCCCAACGAGG	Anemia, nonspherocytic hemolytic, due to G6PD deficiency
137852346	G6PD	NM_000402.4(G6PD):c.896G>A (p.Cys299Tyr)	GATGCTGTRTCTGGTGGCCATGG	Anemia, nonspherocytic hemolytic, due to G6PD deficiency
267606836	G6PD	NM_000402.4(G6PD):c.634C>T (p.Arg212Trp)	CCTGCAGAGCTCTGACYGGCTGT	
398123546	G6PD	NM_000402.4(G6PD):c.1450C>T (p.Arg484Cys)	CCAGATGCACTTCGTGYGCAGGT	Favism, susceptibility to, Anemia, nonspherocytic hemolytic, due to G6PD deficiency, not provided
137852330	G6PD	NM_000402.4(G6PD):c.682C>T (p.Arg228Cys)	CCGTGAGGACCAGATCTACYGCA	
137852334	G6PD	NM_000402.4(G6PD):c.1249C>T (p.Arg417Cys)	CCACCAGCAGTGCAAGYGCAACG	Anemia, nonspherocytic hemolytic, due to G6PD deficiency
137852345	G6PD	NM_000402.4(G6PD):c.1172C>T (p.Ala391Val)	CCTGCGCTGCGGCAAGGYCCTGA	
28937909	GAA	NM_000152.3(GAA):c.1927G>A (p.Gly643Arg)	CTGGTCRGGGCCGACGTCTGCGG	GLYCOGEN STORAGE DISEASE II, ADULT FORM
796051877	GAA	NM_000152.3(GAA):c.1437G>A (p.Lys479=)	GAARGTAGGGCGAGGGTCCAGGG, GGAARGTAGGGCGAGGGTCCAGG	Glycogen storage disease, type II
369532274	GAA	NM_000152.3(GAA):c.2512C>T (p.Gln838Ter)	CCACAGAGTCCCGCYAGCAGCCC	Glycogen storage disease, type II, not provided
121907942	GAA	NM_000152.3(GAA):c.1634C>T (p.Pro545Leu)	CCCACCCTACGTGCYTGGTCAGC	GLYCOGEN STORAGE DISEASE II, ADULT FORM

121907943	GAA	NM_000152.3(GAA):c.2560C>T (p.Arg854Ter)	CCAAGGGTGGGGAGGCCYGAGGG	Glycogen storage disease, type II
587777308	GABRA1	NM_000806.5(GABRA1):c.335G>A (p.Arg112Gln)	CCTCCRGTTAAATAACCTAATGG	Epileptic encephalopathy, early infantile, 19, not specified, not provided
397514737	GABRG2	NM_000816.3(GABRG2):c.968G>A (p.Arg323Gln)	TGCCCRGAAATCGCTCCCCAAGG	Generalized epilepsy with febrile seizures plus 3, not provided
121909673	GABRG2	NM_000816.3(GABRG2):c.245G>A (p.Arg82Gln)	AAACTTCRGCCTGATATAGGAGG	Epilepsy, childhood absence 2, Familial febrile seizures 8, not provided
121909674	GABRG2	NM_198903.2(GABRG2):c.1312C>T (p.Gln438Ter)	CCCAAGATCAGCAACCATTYAAA, CCAAGATCAGCAACCATTYAAAT	Generalized epilepsy with febrile seizures plus 3
796052504	GABRG2	NM_000816.3(GABRG2):c.406C>T (p.Arg136Ter)	CCATTAAAGTCCTCYGATTGAAC	not provided
28940882	GALE	NM_000403.3(GALE):c.269G>A (p.Gly90Glu)	TTTGCGGRGCTCAAGGCCGTGGG, CTTTGCGGRGCTCAAGGCCGTGG	UDPglucose-4-epimerase deficiency
28940885	GALE	NM_000403.3(GALE):c.956G>A (p.Gly319Glu)	TGGRGTGGACAGCAGCCTTAGGG, CTGGRGTGGACAGCAGCCTTAGG	UDPglucose-4-epimerase deficiency, not provided
137853860	GALE	NM_000403.3(GALE):c.715C>T (p.Arg239Trp)	CCCTTCTCTGCAGGTGTCYGGGA, CCTTCTCTGCAGGTGTCYGGGAT	UDPglucose-4-epimerase deficiency
111033608	GALK1	NM_000154.1(GALK1):c.1144C>T (p.Gln382Ter)	CCACCTTCTACCTCTCTYAAGCA, CCTTCTACCTCTCTYAAGCAGCC	Deficiency of galactokinase
118204447	GALNS	NM_000512.4(GALNS):c.178G>A (p.Asp60Asn)	TTGRACCGGATGGCTGCAGAAGG	Mucopolysaccharidosis, MPS-IV-A
398123438	GALNS	NM_000512.4(GALNS):c.463G>A (p.Gly155Arg)	CACRGATTTGATGAGTGGTTTGG, TGAAGCACRGATTTGATGAGTGG	Mucopolysaccharidosis, MPS-IV-A, not provided
118204437	GALNS	NM_000512.4(GALNS):c.1156C>T (p.Arg386Cys)	CCTATCTTCTATTACYGTGGCGA	Mucopolysaccharidosis, MPS-IV-A, not provided
367543255	GALT	NM_000155.3(GALT):c.389G>A (p.Cys130Tyr)	CATGTRCTTCCACCCCTGGTCGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033829	GALT	NM_000155.3(GALT):c.98G>A (p.Arg33His)	TATCCRCTACAACCCGCTGCAGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033675	GALT	NM_000155.3(GALT):c.368G>A (p.Arg123Gln)	CTCRAGGAGTCTGGTAACTATGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033694	GALT	NM_000155.3(GALT):c.443G>A (p.Arg148Gln)	TCCRGGCTGTTGTTGATGCATGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033704	GALT	NM_000155.3(GALT):c.462G>A (p.Trp154Ter)	TGCATGRGCCTCAGTCACAGAGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033723	GALT	NM_000155.3(GALT):c.56	CTGCCAGRTAAGGGTGTCAGGGG	Deficiency of UDPglucose-hexose-1-phosphate

		4+1G>A	, ACTGCCAGRTAAGGGTGTCAGGG	uridylyltransferase
111033736	GALT	NM_000155.3(GALT):c.607G>A (p.Glu203Lys)	GCGTGAGRAGCGATCTCAGCAGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033747	GALT	NM_000155.3(GALT):c.658G>A (p.Glu220Lys)	GCTAATGRAGTACAGCCGCCAGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033784	GALT	NM_000155.3(GALT):c.922G>A (p.Glu308Lys)	GATCARAGGCTGGGGCCAACTGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033802	GALT	NM_000155.3(GALT):c.983G>A (p.Arg328His)	TCCTGCRCTCTGCCACTGTCCGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
367543268	GALT	NM_000155.3(GALT):c.1060-1G>A	CCARGCTGCAGAGAGACTAAGGG, TCCARGCTGCAGAGAGACTAAGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
367543266	GALT	NM_000155.3(GALT):c.961C>T (p.His321Tyr)	CCATTGGCAGCTGCACGCTYATT	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033739	GALT	NM_000155.3(GALT):c.601C>T (p.Arg201Cys)	CCTGCCAGATATTGCCCAGYGTG, CCAGATATTGCCCAGYGTGAGGA	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033774	GALT	NM_000155.3(GALT):c.865C>T (p.Leu289Phe)	CCAAGTATGACAACYTCTTTGAG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033803	GALT	NM_000155.3(GALT):c.986C>T (p.Ser329Phe)	CCCTCCGCTCCTGCGCTYTGCCA, CCTCCGCTCCTGCGCTYTGCCAC	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033804	GALT	NM_000155.3(GALT):c.989C>T (p.Ala330Val)	CCTCCGCTCCTGCGCTCTGYCAC, CCGCTCCTGCGCTCTGYCACTGT	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
367543259	GALT	NM_000155.3(GALT):c.542C>T (p.Ser181Phe)	CCATGATGGGCTGTTYTAACCCC	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
368166217	GALT	NM_000155.3(GALT):c.772C>T (p.Arg258Cys)	CCAGACACTGCTGCTGCCCYGTC	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033649	GALT	NM_000155.3(GALT):c.160C>T (p.Gln54Ter)	CCGCATGAAGCGGCCCTGGYAGG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033686	GALT	NM_000155.3(GALT):c.413C>T (p.Thr138Met)	CCCCTGGTCGGATGTAAYGCTGC, CCCTGGTCGGATGTAAYGCTGCC, CCTGGTCGGATGTAAYGCTGCCA	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
111033845	GALT	NM_000155.3(GALT):c.770C>T (p.Pro257Leu)	CCAGACACTGCTGCTGCYCGTCG	Deficiency of UDPglucose-hexose-1-phosphate uridylyltransferase
121909272	GAMT	NM_000156.5(GAMT):c.506G>A (p.Cys169Tyr)	ACTRCAACCTCACCTCCTGGGGG, TACTRCAACCTCACCTCCTGGGG, CTACTRCAACCTCACCTCCTGGG, CCTACTRCAACCTCACCTCCTGG	Deficiency of guanidinoacetate methyltransferase
80338735	GAMT	NM_000156.5(GAMT):c.32	CAARGTGCCCCTCTGCCCGCAGG	Deficiency of guanidinoacetate methyltransferase, not

		7G>A (p.Lys109=)		provided
119485089	GAN	NM_022041.3(GAN):c.1447C>T (p.Gln483Ter)	CCGAAGTCGTGAGGACGCCYAGG	Giant axonal neuropathy
104894809	GATA1	NM_002049.3(GATA1):c.647G>A (p.Arg216Gln)	CTGTGGCRGAGGGACAGGACAGG	Thrombocytopenia, platelet dysfunction, hemolysis, and imbalanced globin synthesis
387907207	GATA1	NM_002049.3(GATA1):c.646C>T (p.Arg216Trp)	CCACTCCACTGTGGYGGAGGGAC	Thrombocytopenia, platelet dysfunction, hemolysis, and imbalanced globin synthesis
387906630	GATA2	NM_001145661.1(GATA2):c.761C>T (p.Pro254Leu)	CCTACCCCTCCTATGTGCYGGCG, CCCCTCCTATGTGCYGGCGGCTG	Dendritic cell, monocyte, B lymphocyte, and natural killer lymphocyte deficiency
387906632	GATA2	NM_001145661.1(GATA2):c.1009C>T (p.Arg337Ter)	CCACTCATCAAGCCCAAGYGAAG	Lymphedema, primary, with myelodysplasia
104894162	GATA3	NM_001002295.1(GATA3):c.829C>T (p.Arg277Ter)	CCCCACTGTGGCGGYGAGATGGC	Barakat syndrome
56208331	GATA4	NM_002052.4(GATA4):c.1273G>A (p.Asp425Asn)	GCAGRACTCTTGGAACAGCCTGG	Multiple congenital anomalies, Tetralogy of Fallot, Atrial septal defect 2
104894074	GATA4	NM_002052.4(GATA4):c.155C>T (p.Ser52Phe)	CCTCCGTGCTGGGCCTGTYCTAC, CCGTGCTGGGCCTGTYCTACCTC	Atrial septal defect 2
115372595	GATA4	NM_002052.4(GATA4):c.1037C>T (p.Ala346Val)	CCTCCCGCCAGCGGTGYTTCCAG	Atrioventricular septal defect 4
387906769	GATA4	NM_002052.4(GATA4):c.487C>T (p.Pro163Ser)	CCTACTCCAGCCCCTACYCGGCT	Tetralogy of Fallot, Ventricular septal defect 1, Atrioventricular septal defect 4
387906771	GATA4	NM_002052.4(GATA4):c.839C>T (p.Thr280Met)	CCAGACCACCACCACCAYGCTGT	Atrial septal defect 2
387906819	GATA6	NM_005257.5(GATA6):c.1367G>A (p.Arg456His)	TTATGGCRCAGAAACGCCGAGGG, CTTATGGCRCAGAAACGCCGAGG	Pancreatic agenesis and congenital heart disease
80356772	GBA	NM_000157.3(GBA):c.1505G>A (p.Arg502His)	AACCRGTGAGGGCAATGGTGAGG	Gaucher disease
121908311	GBA	NM_000157.3(GBA):c.1246G>A (p.Gly416Ser)	ATGTGGTCRGCTGGACCGACTGG	Gaucher disease, Subacute neuronopathic Gaucher disease, Gaucher disease, type 1
121908298	GBA	NM_001005741.2(GBA):c.983C>T (p.Pro328Leu)	CCAACGCTTGCTGCTGCYCCACT	Gaucher disease, type 1
398123532	GBA	NM_001005741.2(GBA):c.625C>T (p.Arg209Cys)	CCCTGCAGTTGGCCCAGYGTCCC, CCTGCAGTTGGCCCAGYGTCCCG	Gaucher disease, type 1, not provided
398123015	GBA2	NM_020944.2(GBA2):c.2618G>A (p.Arg873His)	TCCRCTCACTGGCCTACATGCGG	
398123013	GBA2	NM_020944.2(GBA2):c.700C>T (p.Arg234Ter)	CCATGCCCTCTATCCCYGAGCCT	
80338673	GBE1	NM_000158.3(GBE1):c.157	TGATTCRACTCATTACGCATGGG,	Glycogen storage disease, type IV, GLYCOGEN STORAGE

		1G>A (p.Arg524Gln)	ATGATTCRACTCATTACGCATGG	DISEASE IV, COMBINED HEPATIC AND MYOPATHIC
786205862	GCDH	NM_000159.3(GCDH):c.675G>A (p.Trp225Ter)	GTGTGRGCTCGGTGTGAAGATGG	Glutaric aciduria, type 1
147611168	GCDH	NM_000159.3(GCDH):c.1240G>A (p.Glu414Lys)	ACACCTACRAAGGTAGGAGCTGG	Glutaric aciduria, type 1, not provided
104894438	GCH1	NM_000161.2(GCH1):c.602G>A (p.Gly201Glu)	GCCTGCTGRAGTCGGGGTAGTGG	Dystonia 5, Dopa-responsive type
104894443	GCH1	NM_000161.2(GCH1):c.633G>A (p.Met211Ile)	CACATRTGTATGGTAATGCGAGG	GTP cyclohydrolase I deficiency
104894444	GCH1	NM_000161.2(GCH1):c.142C>T (p.Gln48Ter)	CCCGAGGCCAAGAGCGCGYAGCC, CCGAGGCCAAGAGCGCGYAGCCC	Dystonia 5, Dopa-responsive type
193922289	GCK	NM_000162.3(GCK):c.214G>A (p.Gly72Arg)	AGTCRGGGACTTCCTCTCCCTGG	Maturity-onset diabetes of the young, type 2
104894008	GCK	NM_000162.3(GCK):c.781G>A (p.Gly261Arg)	CTTCRGGGACTCCGGCGAGCTGG	Maturity-onset diabetes of the young, type 2
104894012	GCK	NM_000162.3(GCK):c.1363G>A (p.Val455Met)	CTCGGCGRTGGCCTGTAAGAAGG	Hyperinsulinemic hypoglycemia familial 3
104894016	GCK	NM_000162.3(GCK):c.1132G>A (p.Ala378Thr)	GCGCRCTGCGCACATGTGCTCGG	Maturity-onset diabetes of the young, type 2
397514580	GCK	NM_000162.3(GCK):c.1015G>A (p.Glu339Lys)	GGTGRAGAGGTGTGCGGAGGAGG, GCAGGTGRAGAGGTGTGCGGAGG	Maturity-onset diabetes of the young, type 2
587780347	GCK	NM_000162.3(GCK):c.706G>A (p.Glu236Lys)	CATGRAGGAGATGCAGAATGTGG	Diabetes mellitus, gestational
104894014	GCK	NM_000162.3(GCK):c.1367C>T (p.Ala456Val)	CCCTGGTCTCGGCGGTGGYCTGT, CCTGGTCTCGGCGGTGGYCTGTA	Hyperinsulinemic hypoglycemia familial 3
80356655	GCK	NM_000162.3(GCK):c.683C>T (p.Thr228Met)	CCGACCTCCACCCCAGGCAYGGG, CCTCCACCCCAGGCAYGGGCTGC	Permanent neonatal diabetes mellitus, Maturity-onset diabetes of the young, type 2
56141211	GCNT2	NM_001491.2(GCNT2):c.1043G>A (p.Gly348Glu)	ATGRAAACGGAGACTTAAAGTGG	I blood group system
137853339	GCNT2	NM_145649.4(GCNT2):c.505G>A (p.Ala169Thr)	CCAGRCTGACCTGAACTGCCTGG	I blood group system
397515432	GDAP1	NM_018972.2(GDAP1):c.980G>A (p.Gly327Asp)	GTTGRTTTGCTTGCAGGAGTGGG, GGTTGRTTTGCTTGCAGGAGTGG	Charcot-Marie-Tooth disease, recessive intermediate A
387906946	GDF3	NM_020634.1(GDF3):c.820	CCAGCTATTCATTAACTTCYGGG	Microphthalmia, isolated, with coloboma 6

		C>T (p.Arg274Trp)		
36119840	GDNF	NM_000514.3(GDNF):c.277C>T (p.Arg93Trp)	TGCCRATTCCGCTCTCTTCTAGG	Congenital central hypoventilation, Hirschsprung disease 3, not specified
RS# (dbSNP)	GeneSymbol	Name	gRNAall	Phenotypes
58064122	GFAP	NM_002055.4(GFAP):c.715C>T (p.Arg239Cys)	CCCTGAAAGAGATCYGCACGCAG	Alexander disease, not provided
121908192	GFER	NM_005262.2(GFER):c.581G>A (p.Arg194His)	ATGAGCRCTGGCGCGACGGCTGG	Myopathy, mitochondrial progressive, with congenital cataract, hearing loss, and developmental delay, not provided
119470019	GFM1	NM_024996.5(GFM1):c.139C>T (p.Arg47Ter)	CCTAATGAAAAAATAYGAAATAT	Combined oxidative phosphorylation deficiency 1
121909678	GGCX	NM_000821.6(GGCX):c.1672G>A (p.Gly558Arg)	GCTGCAGRGGGAAGTGACTGTGG	Pseudoxanthoma elasticum-like disorder with multiple coagulation factor deficiency
121909682	GGCX	NM_000821.6(GGCX):c.1427G>A (p.Arg476His)	TGACCRCTTCCAGCAGAGGTGGG, ATGACCRCTTCCAGCAGAGGTGG	Pseudoxanthoma elasticum-like disorder with multiple coagulation factor deficiency
121909683	GGCX	NM_000821.6(GGCX):c.763G>A (p.Val255Met)	TGCTGGTCRTGCACTGGGGTGGG	Pseudoxanthoma elasticum-like disorder with multiple coagulation factor deficiency
121909680	GGCX	NM_000821.6(GGCX):c.1120C>T (p.Gln374Ter)	CCTGCTCTACCTCCTGGAGYAGC	Pseudoxanthoma elasticum-like disorder with multiple coagulation factor deficiency
121909684	GGCX	NM_000821.6(GGCX):c.899C>T (p.Ser300Phe)	CCTAGGTATGTTCTYCTACGTCA	Pseudoxanthoma elasticum-like disorder with multiple coagulation factor deficiency
387907000	GIPC3	NM_133261.2(GIPC3):c.903G>A (p.Trp301Ter)	AGTGTGRGCCGCCATCGGCGAGG	Deafness, autosomal recessive 15
387907002	GIPC3	NM_133261.2(GIPC3):c.565C>T (p.Arg189Cys)	CCCAGCCCTTCACCCTGYGCCTG, CCAGCCCTTCACCCTGYGCCTGG	Deafness, autosomal recessive 15
104893963	GJA1	NM_000165.4(GJA1):c.61G>A (p.Gly21Arg)	CAACTGCTRGAGGGAAGGTGTGG	Oculodentodigital dysplasia
104893965	GJA1	NM_000165.4(GJA1):c.1127G>A (p.Arg376Gln)	CAGACCTCRGCCTGATGACCTGG	Hypoplastic left heart syndrome, Atrioventricular septal defect and common atrioventricular junction
28931600	GJA1	NM_000165.4(GJA1):c.427G>A (p.Gly143Ser)	CATRGTAAGGTGAAAATGCGAGG	Syndactyly type 3
387906616	GJA1	NM_000165.4(GJA1):c.31C>T (p.Leu11Phe)	CCTTAGGCAAACTCYTTGACAAG	Oculodentodigital dysplasia
397514703	GJA3	NM_021954.3(GJA3):c.5G>A (p.Gly2Asp)	ATGGRCGACTGGAGCTTTCTGGG, AATGGRCGACTGGAGCTTTCTGG	Zonular pulverulent cataract 3
398122937	GJA3	NM_021954.3(GJA3):c.427G>A (p.Gly143Arg)	GCATGGCCRGGGCGCTGCTGCGG	Zonular pulverulent cataract 3
121917825	GJA3	NM_021954.3(GJA3):c.560	CCGCTGGCCCTGCCYCAACACGG	Zonular pulverulent cataract 3

		C>T (p.Pro187Leu)		
387906612	GJA5	NM_005266.6(GJA5):c.145 C>T (p.Gln49Ter)	CCTGGGGGGATGAGYAGGCTGAT	Atrial fibrillation, familial, 11
397515627	GJA8	NM_005267.4(GJA8):c.566 C>T (p.Pro189Leu)	CCGGTGGCCCTGCCYCAATGTGG	Cataract 1
786204123	GJB1	NM_000166.5(GJB1):c.425 G>A (p.Arg142Gln)	GGTGTTCCRGCTGTTGTTTGAGG	Charcot-Marie-Tooth Neuropathy X
104894814	GJB1	NM_001097642.2(GJB1):c. 658C>T (p.Arg220Ter)	CCGGGCCTGTGCCCGCYGAGCCC	X-linked hereditary motor and sensory neuropathy
104894824	GJB1	NM_000166.5(GJB1):c.164 C>T (p.Thr55Ile)	CCTTCATCTGCAACAYACTCCAG	X-linked hereditary motor and sensory neuropathy
587777876	GJB1	NM_000166.5(GJB1):c.77C >T (p.Ser26Leu)	CCGAGTATGGCTCTYGGTCATCT	X-linked hereditary motor and sensory neuropathy
587777879	GJB1	NM_000166.5(GJB1):c.790 C>T (p.Arg264Cys)	CCCTGAAAGACATACTGYGCCGC , CCTGAAAGACATACTGYGCCGCA	X-linked hereditary motor and sensory neuropathy
587781246	GJB1	NM_000166.5(GJB1):c.688 C>T (p.Arg230Cys)	CCGCTCCAATCCACCTTCCYGCA, CCAATCCACCTTCCYGCAAGGGC	Charcot-Marie-Tooth disease
116840819	GJB1	NM_000166.5(GJB1):c.223 C>T (p.Arg75Trp)	CCCCATCTCCCATGTGYGGCTGT, CCCATCTCCCATGTGYGGCTGTG, CCATCTCCCATGTGYGGCTGTGG	X-linked hereditary motor and sensory neuropathy
587783645	GJB2	NM_004004.5(GJB2):c.158 G>A (p.Cys53Tyr)	GTCTRCAACACCCTGCAGCCAGG	Hearing impairment
80338940	GJB2	NM_004004.5(GJB2):c.-23+1G>A	GCAGRTGAGCCCGCCGGCCCCGG	Deafness, autosomal recessive 1A, Hearing impairment
104894402	GJB2	NM_004004.5(GJB2):c.223 C>T (p.Arg75Trp)	CCCCATCTCCCACATCYGGCTAT, CCCATCTCCCACATCYGGCTATG, CCATCTCCCACATCYGGCTATGG	Deafness, autosomal dominant 3a
76434661	GJB2	NM_004004.5(GJB2):c.416 G>A (p.Ser139Asn)	CCCGGAAGAAGATGYTGCTTGTG	Deafness, autosomal recessive 1A, Hearing impairment
72555392	GLB1	NM_000404.2(GLB1):c.176 G>A (p.Arg59His)	CCCRTGTGCCCCGCTTCTACTGG	Mucopolysaccharidosis, MPS-IV-B, Infantile GM1 gangliosidosis, Juvenile GM>1< gangliosidosis, Gangliosidosis GM1 type 3, GM1-GANGLIOSIDOSIS, TYPE I, WITH CARDIAC INVOLVEMENT, not provided
398123351	GLB1	NM_000404.2(GLB1):c.176 9G>A (p.Arg590His)	GGCCRCTATTGGCCAGCCCGGGG , TGGCCRCTATTGGCCAGCCCGGG, TTGGCCRCTATTGGCCAGCCCGG	Mucopolysaccharidosis, MPS-IV-B, Infantile GM1 gangliosidosis, Juvenile GM>1< gangliosidosis, Gangliosidosis GM1 type 3, not provided
398123353	GLB1	NM_000404.2(GLB1):c.397	ACTARGGAGGATTACCTGCTTGG	Mucopolysaccharidosis, MPS-IV-B, Infantile GM1

		-1G>A		gangliosidosis, Juvenile GM>1< gangliosidosis, Gangliosidosis GM1 type 3, not provided
72555359	GLB1	NM_000404.2(GLB1):c.1369C>T (p.Arg457Ter)	CCCCCAGGGAGTCCTTGAGYGAA, CCCCAGGGAGTCCTTGAGYGAAA, CCCAGGGAGTCCTTGAGYGAAAC, CCAGGGAGTCCTTGAGYGAAACA	Infantile GM1 gangliosidosis
72555366	GLB1	NM_000404.2(GLB1):c.622C>T (p.Arg208Cys)	CCTGCAGAAGCGCTTTYGCCACC	Mucopolysaccharidosis, MPS-IV-B, Infantile GM1 gangliosidosis, Juvenile GM>1< gangliosidosis, Gangliosidosis GM1 type 3, not provided
72555370	GLB1	NM_000404.2(GLB1):c.202C>T (p.Arg68Trp)	CCGCTTCTACTGGAAGGACYGGC	Juvenile GM>1< gangliosidosis
121964980	GLDC	NM_000170.2(GLDC):c.2216G>A (p.Arg739His)	TCTGTCRCCCTGGAGACTTCGGG, ATCTGTCRCCCTGGAGACTTCGG	Non-ketotic hyperglycinemia
121964977	GLDC	NM_000170.2(GLDC):c.2405C>T (p.Ala802Val)	CCTGTGGGAACCGTCAGTGYGGC	Non-ketotic hyperglycinemia
121917707	GLI2	NM_005270.4(GLI2):c.1323G>A (p.Trp441Ter)	CCACTGRGAAGACTGCACCAAGG	Holoprosencephaly 9
116840748	GLI3	NM_000168.5(GLI3):c.2110C>T (p.Gln704Ter)	CCCCAACAGACATCTYAGCCAAG, CCCAACAGACATCTYAGCCAAGC	Pallister-Hall syndrome
116840770	GLI3	NM_000168.5(GLI3):c.3481C>T (p.Gln1161Ter)	CCGACCTGCCCATTYAGTGGAAC	Pallister-Hall syndrome
281864919	GLRA1	NM_000171.3(GLRA1):c.1259G>A (p.Arg420His)	ATCCCRCATTGGCTTCCCCATGG	Hyperekplexia hereditary
116474260	GLRA1	NM_001146040.1(GLRA1):c.1132G>A (p.Gly378Ser)	CCTGTAGACAGGCTGGGCYCATC	Hyperekplexia hereditary
121909749	GLRB	NM_000824.4(GLRB):c.752G>A (p.Gly251Asp)	GAAAGRCTACTACACATGCGTGG	Hyperekplexia 2
121909736	GLUD1	NM_005271.3(GLUD1):c.953G>A (p.Arg318Lys)	ATGARATATTTACATCGTTTTGG	Hyperinsulinism-hyperammonemia syndrome
397509422	GMPPB	NM_013334.3(GMPPB):c.1081G>A (p.Asp361Asn)	CCGTTGAGGTAGAGCTCATYATT	Limb-girdle muscular dystrophy-dystroglycanopathy, type C14, Congenital muscular dystrophy-dystroglycanopathy with brain and eye anomalies, type A14
397509423	GMPPB	NM_013334.3(GMPPB):c.220C>T (p.Arg74Ter)	CCTTCCAGCTGGGAATCYGAATC	Congenital muscular dystrophy-dystroglycanopathy with brain and eye anomalies, type A14
397509424	GMPPB	NM_013334.3(GMPPB):c.6	CCGCTGACGCTGAGCACCYCGAA	Limb-girdle muscular dystrophy-dystroglycanopathy, type

		4C>T (p.Pro22Ser)		C14
397509425	GMPPB	NM_021971.2(GMPPB):c.553C>T (p.Arg185Cys)	CCCTGCAGTGCTGCAGYGCATCC, CCTGCAGTGCTGCAGYGCATCCA	Limb-girdle muscular dystrophy-dystroglycanopathy, type C14, Congenital muscular dystrophy-dystroglycanopathy with brain and eye anomalies, type A14, Congenital muscular dystrophy-dystroglycanopathy with mental retardation, type B14
202160208	GMPPB	NM_013334.3(GMPPB):c.860G>A (p.Arg287Gln)	CCGCAGCACCGTGCACCGCYGGA	Congenital muscular dystrophy-dystroglycanopathy with mental retardation, type B14
137853227	GNAI2	NM_002070.3(GNAI2):c.536G>A (p.Arg179His)	GGACCCRCGTAAAGACCACGGGG, CGGACCCRCGTAAAGACCACGGG, ACGGACCCRCGTAAAGACCACGG	Granulosa cell tumor of the ovary
398122923	GNAL	NM_001142339.2(GNAL):c.409G>A (p.Val137Met)	TGACCATRTGAAAAAACTTTGGG, TTGACCATRTGAAAAAACTTTGG	Dystonia 25
397514698	GNAQ	NM_002072.4(GNAQ):c.548G>A (p.Arg183Gln)	GAGTTCRAGTCCCCACCACAGGG, AGAGTTCRAGTCCCCACCACAGG	Sturge-Weber syndrome, Capillary malformations, congenital, 1
121913495	GNAS	NM_000516.5(GNAS):c.602G>A (p.Arg201His)	CGCTGCCDTGTCCTGACTTCTGG	McCune-Albright syndrome, Somatotroph adenoma, Sex cord-stromal tumor, Cushing syndrome
137854539	GNAS	NM_001077488.3(GNAS):c.347C>T (p.Pro116Leu)	CCATGAGCAACCTGGTGCYCCCC	Pseudohypoparathyroidism type 1A, Pseudopseudohypoparathyroidism
121908625	GNE	NM_001128227.2(GNE):c.1820G>A (p.Gly607Glu)	GATGRGCCTGATTGTTCCTGTGG	Inclusion body myopathy 2
62541771	GNE	NM_001128227.2(GNE):c.1985C>T (p.Ala662Val)	GTTTCRCAGCTTGGATGAGATGG	Inclusion body myopathy 2, Nonaka myopathy
121434440	GNPAT	NM_014236.3(GNPAT):c.631C>T (p.Arg211Cys)	CCTTTTTCATGCGGYGTACCTTT	Rhizomelic chondrodysplasia punctata type 2
137852885	GNPTG	NM_032520.4(GNPTG):c.316G>A (p.Gly106Ser)	TCCTCRGGTGAGTGGGGCCGGGG, ATCCTCRGGTGAGTGGGGCCGGG, GATCCTCRGGTGAGTGGGGCCGG	Mucolipidosis III Gamma
193302848	GNPTG	NM_032520.4(GNPTG):c.196C>T (p.Arg66Ter)	CCCGTGCATCTCTTCYGACTCTC, CCGTGCATCTCTTCYGACTCTCG	Mucolipidosis III Gamma
193302854	GNPTG	NM_032520.4(GNPTG):c.610-1G>T	CCTGCATCCTCCACCTTCAYGGC	Mucolipidosis III Gamma
104893842	GNRHR	NM_000406.2(GNRHR):c.416G>A (p.Arg139His)	ACCRCTCCCTGGCTATCACGAGG	

104893847	GNRHR	NM_000406.2(GNRHR):c.959C>T (p.Pro320Leu)	CCCATGCTTTGATCYACTTATCT	
267606849	GP1BA	NM_000173.6(GP1BA):c.1620G>A (p.Trp540Ter)	CTGRCTGCTCTTTGCCTCTGTGG	Bernard-Soulier syndrome, type A1
121908063	GP1BA	NM_000173.6(GP1BA):c.217C>T (p.Leu73Phe)	CCTGATGCCTTACACTCGCYTCA	Bernard-Soulier syndrome, type A2, autosomal dominant
137853582	GPI	NM_000175.3(GPI):c.475G>A (p.Gly159Ser)	TTGGCRGCTCCGACCTGGTGAGG	Hemolytic anemia, nonspherocytic, due to glucose phosphate isomerase deficiency
137853585	GPI	NM_000175.3(GPI):c.1615G>A (p.Asp539Asn)	CTCACRACGCTTCTACCAATGGG, TCTCACRACGCTTCTACCAATGG	Hemolytic anemia, nonspherocytic, due to glucose phosphate isomerase deficiency
61754634	GPI	NM_000175.3(GPI):c.671C>T (p.Thr224Met)	CCATCACGAATGCAGAGAYGGCG	Hemolytic anemia, nonspherocytic, due to glucose phosphate isomerase deficiency
267606852	GPI	NM_000175.3(GPI):c.14C>T (p.Thr5Ile)	CCCGCCATGGCCGCTCTCAYCCG, CCGCCATGGCCGCTCTCAYCCGG, CCATGGCCGCTCTCAYCCGGGAC	Hemolytic anemia, nonspherocytic, due to glucose phosphate isomerase deficiency
58933950	GPR143	NM_000273.2(GPR143):c.455G>A (p.Ser152Asn)	CCATTTCCTCGGTGAATACYTCA	Ocular albinism, type I, not provided
387907138	GPR179	NM_001004334.3(GPR179):c.598C>T (p.Arg200Ter)	CCCCTGCCCTGAAGAAGYGAGTG, CCCTGCCCTGAAGAAGYGAGTGT, CCTGCCCTGAAGAAGYGAGTGTT	Congenital stationary night blindness, type 1E
267606854	GPSM2	NM_013296.4(GPSM2):c.379C>T (p.Arg127Ter)	CCATAGTTTGTTGTCAGYGACAC	Chudley-McCullough syndrome
769967246	GPX4	NM_001039848.2(GPX4):c.381C>A (p.Tyr127Ter)	CCTGCACGCCCGATAYGCTGAGT	Spondylometaphyseal dysplasia Sedaghatian type
180177312	GRHPR	NM_012203.1(GRHPR):c.478G>A (p.Gly160Arg)	ATCATCRGGCTGGGGCGCATAGG	Primary hyperoxaluria, type II
180177314	GRHPR	NM_012203.1(GRHPR):c.494G>A (p.Gly165Asp)	CTCTAGRCCAGGCCATTGCTCGG	Primary hyperoxaluria, type II
180177322	GRHPR	NM_012203.1(GRHPR):c.904C>T (p.Arg302Cys)	CCACCCACAGAACCYGCAACACC	Primary hyperoxaluria, type II
137852350	GRIA3	NM_007325.4(GRIA3):c.2497G>A (p.Gly833Arg)	ACTTGTCRGAGGTCTGGGGCTGG	Mental retardation, X-linked, syndromic, wu type
397518470	GRIN2A	NM_000833.4(GRIN2A):c.1553G>A (p.Arg518His)	TGAGGAACRTTCTGAAGTGGTGG	Focal epilepsy with speech disorder with or without mental retardation
796052571	GRIN2B	NM_000834.3(GRIN2B):c.1858G>A (p.Val620Met)	TACCTRTGCAGAACCCAAAGGGG, GTACCTRTGCAGAACCCAAAGGG	not provided

			, CGTACCTRTGCAGAACCCAAAGG	
397514556	GRIN2B	NM_000834.3(GRIN2B):c.1658C>T (p.Pro553Leu)	CCCTTCCTCAGAGCYATTCAGCG	Mental retardation, autosomal dominant 6
63750331	GRN	NM_002087.3(GRN):c.3G>A (p.Met1Ile)	CATRTGGACCCTGGTGAGCTGGG, CCATRTGGACCCTGGTGAGCTGG	Frontotemporal dementia, ubiquitin-positive, not provided
606231221	GRN	NM_002087.3(GRN):c.835+1G>A	GCACACAGRTACCAGAGGCAGGG	Frontotemporal dementia, ubiquitin-positive
63750077	GRN	NM_002087.3(GRN):c.373C>T (p.Gln125Ter)	CCGTGGGTGCCATCYAGTGCCCT	Frontotemporal dementia, ubiquitin-positive, not provided
63751294	GRN	NM_002087.3(GRN):c.1477C>T (p.Arg493Ter)	CCTGCAACGTGAAGGCTYGATCC	Frontotemporal dementia, ubiquitin-positive, not provided
193026789	GRN	NM_002087.3(GRN):c.1212C>A (p.Cys404Ter)	CCACCAGCACTGCTGYCCCCAGG	Frontotemporal dementia
587777289	GSC	NM_173849.2(GSC):c.400C>T (p.Gln134Ter)	CAGCATCTRGTGCGGTACCGGGG	Short stature, auditory canal atresia, mandibular hypoplasia, skeletal abnormalities
121909307	GSS	NM_000178.2(GSS):c.491G>A (p.Arg164Gln)	TGTGCACCRGTGGGTCCCCTGGG	Gluthathione synthetase deficiency
121909124	GUCA1B	NM_002098.5(GUCA1B):c.469G>A (p.Gly157Arg)	GAGAATRGAGATGGTAAGAGGGG, TGAGAATRGAGATGGTAAGAGGG, ATGAGAATRGAGATGGTAAGAGG	Retinitis pigmentosa 48, not provided
61750173	GUCY2D	NM_000180.3(GUCY2D):c.2513G>A (p.Arg838His)	CCGGGAGCRCACGGAGGAGCTGG	Cone-rod dystrophy 6, not provided
121918179	GUSB	NM_000181.3(GUSB):c.1521G>A (p.Trp507Ter)	ACTCTTGRTATCACGACTACGGG, TACTCTTGRTATCACGACTACGG	Mucopolysaccharidosis type VII
398123234	GUSB	NM_000181.3(GUSB):c.1084G>A (p.Asp362Asn)	GGGCTTCRACTGGCCGCTGCTGG	Mucopolysaccharidosis type VII, not provided
377519272	GUSB	NM_000181.3(GUSB):c.1616_1653del38 (p.Ser539Argfs*8)	CCATACTCRCTCTGAATAATGGG	Mucopolysaccharidosis type VII
587779400	GUSB	NM_000181.3(GUSB):c.530C>T (p.Thr177Ile)	CCATCAACAACACACTCAYCCCC	Mucopolysaccharidosis type VII
121918181	GUSB	NM_000181.3(GUSB):c.526C>T (p.Leu176Phe)	CCATCAACAACACAYTCACCCCC	Mucopolysaccharidosis type VII, not provided
121434584	GYS1	NM_002103.4(GYS1):c.1384C>T (p.Arg462Ter)	CCTGACCACCATCCGCYGAATCG	Glycogen storage disease 0, muscle

121918419	GYS2	NM_021957.3(GYS2):c.736C>T (p.Arg246Ter)	CCACCGGTACTGCATGGAGYGAG, CCGGTACTGCATGGAGYGAGCTT	Hypoglycemia with deficiency of glycogen synthetase in the liver
137853101	HADH	NM_005327.4(HADH):c.118G>A (p.Ala40Thr)	GATGGGCRCCGGCATTGCCCAGG	Deficiency of 3-hydroxyacyl-CoA dehydrogenase
137853103	HADH	NM_005327.4(HADH):c.773C>T (p.Pro258Leu)	CCGGTTACCCCATGGGCCYATTT	Hyperinsulinemic hypoglycemia, familial, 4
121913134	HADHB	NM_000183.2(HADHB):c.1331G>A (p.Arg444Lys)	AACARATTACGGAAAGAAGGAGG, GCCAACARATTACGGAAAGAAGG	Mitochondrial trifunctional protein deficiency
104894695	HAMP	NM_021175.3(HAMP):c.166C>T (p.Arg56Ter)	CCCATGTTCCAGAGGYGAAGGAG, CCATGTTCCAGAGGYGAAGGAGG	Hemochromatosis type 2B
74315322	HAX1	NM_006118.3(HAX1):c.568C>T (p.Gln190Ter)	CCAGATCTTGATTCCYAGGTTTC	Severe congenital neutropenia 3, autosomal recessive
41417548	HBA2	NM_000517.4(HBA2):c.314G>A (p.Cys105Tyr)	CCACTRCCTGCTGGTGACCCTGG	Hemoglobin H disease, nondeletional
63750783	HBB	NM_000518.4(HBB):c.47G>A (p.Trp16Ter)	CCTGTRGGGCAAGGTGAACGTGG	beta^0^ Thalassemia
34999973	HBB	NM_000518.4(HBB):c.-140C>T	CCTCACCCTGTGGAGCCAYACCC	beta Thalassemia
34883338	HBB	NM_000518.4(HBB):c.-50-92C>T	CCTCACCCTGTGGAGCYACACCC	
35378915	HBG1	NM_000559.2(HBG1):c.-170G>A	CTTRACCAATAGCCTTGACAAGG	Fetal hemoglobin quantitative trait locus 1
35983258	HBG1	NM_000559.2(HBG1):c.-53-196C>T	CCTCTTGGGGGCCCCTTCYCCAC	Fetal hemoglobin quantitative trait locus 1
281860601	HBG1	NM_000559.2(HBG1):c.-167C>T	CCAGCCTTGCCTTGACYAATAGC	Fetal hemoglobin quantitative trait locus 1
34474104	HBG2	NM_000184.2(HBG2):c.190C>T (p.His64Tyr)	CCCCAAAGTCAAGGCAYATGGCA, CCCAAAGTCAAGGCAYATGGCAA, CCAAAGTCAAGGCAYATGGCAAG	Cyanosis, transient neonatal
35103459	HBG2	NM_000184.2(HBG2):c.277C>T (p.His93Tyr)	CCCAGCTGAGTGAACTGYACTGT, CCAGCTGAGTGAACTGYACTGTG	Cyanosis, transient neonatal

587776864	HBG2	NM_000184.2(HBG2):c.202G>A (p.Val68Met)	CCCAAGGAAGTCAGCAYCTTCTT, CCAAGGAAGTCAGCAYCTTCTTG	Cyanosis, transient neonatal
193929392	HCCS	NM_005333.4(HCCS):c.475G>A (p.Glu159Lys)	GAATAACRAGCAGGCTTGGAAGG	Microphthalmia, syndromic, 7
318240758	HCFC1	NM_005334.2(HCFC1):c.674G>A (p.Ser225Asn)	GATGARTGGCTGCAGGCTGGGGG, GGATGARTGGCTGCAGGCTGGGG, GGGATGARTGGCTGCAGGCTGGG, CGGGATGARTGGCTGCAGGCTGG	Mental retardation 3, X-linked, not provided
397515486	HCFC1	NM_005334.2(HCFC1):c.218C>T (p.Ala73Val)	CCAGTGGTTCATCCCAGYCGTGA	Mental retardation 3, X-linked
398122909	HDAC8	NM_018486.2(HDAC8):c.958G>A (p.Gly320Arg)	ACTTGACCRGGGTCATCCTAGGG	Cornelia de Lange syndrome 5
387907052	HEPACAM	NM_152722.4(HEPACAM):c.292C>T (p.Arg98Cys)	CCTGACTATCGAGACYGTATCCG	Megalencephalic leukoencephalopathy with subcortical cysts 2a
104893742	HESX1	NM_003865.2(HESX1):c.445G>A (p.Glu149Lys)	TCTAGAGRAAGACAGAATCCAGG	Growth hormone deficiency with pituitary anomalies
121907954	HEXA	NM_000520.4(HEXA):c.805G>A (p.Gly269Ser)	ACCARGTAAGAATGATGTCTGGG, GACCARGTAAGAATGATGTCTGG	Tay-Sachs disease, Gm2-gangliosidosis, adult
121907957	HEXA	NM_000520.4(HEXA):c.509G>A (p.Arg170Gln)	TCCTCACCRGGGCTTGCTGTTGG	Tay-Sachs disease
121907980	HEXA	NM_000520.4(HEXA):c.805+1G>A	ACCAGRTAAGAATGATGTCTGGG, GACCAGRTAAGAATGATGTCTGG	
1800429	HEXA	NM_000520.4(HEXA):c.598G>A (p.Val200Met)	GAACRTGTTCCACTGGCATCTGG	Tay-Sachs disease, B1 variant
121907966	HEXA	NM_000520.4(HEXA):c.1495C>T (p.Arg499Cys)	CCTGACATTTGCCTATGAAYGTT	Tay-Sachs disease, Gm2-gangliosidosis, adult-onset
121907972	HEXA	NM_000520.4(HEXA):c.508C>T (p.Arg170Trp)	CCCCGCTTTCCTCACYGGGGCTT, CCCGCTTTCCTCACYGGGGCTTG	Tay-Sachs disease
76173977	HEXA	NM_000520.4(HEXA):c.1073+1G>A	CCCTCCTTCCTTCCTCAYGTCTG, CCTCCTTCCTTCCTCAYGTCTGG	Tay-Sachs disease, not provided
770932296	HEXA	NM_000520.4(HEXA):c.806-7G>A	CCAGGGATACCTAAGCYAAGAGA	Tay-Sachs disease
121907983	HEXB	NM_000521.3(HEXB):c.1514G>A (p.Arg505Gln)	AGGCCTCRGGCAAGTGCTGTTGG	Sandhoff disease, adult type

121907986	HEXB	NM_000521.3(HEXB):c.850C>T (p.Arg284Ter)	CCAGATTACGAGGAATTYGAGTC	Sandhoff disease, Sandhoff disease, infantile
1800562	HFE	NM_000410.3(HFE):c.845G>A (p.Cys282Tyr)	ACGTRCCAGGTGGAGCACCCAGG	Hemochromatosis type 1, Microvascular complications of diabetes 7, Transferrin serum level quantitative trait locus 2, not specified, not provided
587777269	HFM1	NM_001017975.4(HFM1):c.2206G>A (p.Gly736Ser)	CCAAGTCAAAGTAACAAACYATA	Premature ovarian failure 9
397515347	HGD	NM_000187.3(HGD):c.16-1G>A	TACARTACATTTCTGGATTTGGG, CTACARTACATTTCTGGATTTGG	Alkaptonuria
28942100	HGD	NM_000187.3(HGD):c.688C>T (p.Pro230Ser)	CCTCGTGATTTCTTGATAYCCAT	Alkaptonuria
398124544	HGSNAT	NM_152419.2(HGSNAT):c.1250+1G>A	CTACRTAAGCGAACCCCTGGGGG, CCTACRTAAGCGAACCCCTGGGG, CCCTACRTAAGCGAACCCCTGGG, GCCCTACRTAAGCGAACCCCTGG	not provided
112029032	HGSNAT	NM_152419.2(HGSNAT):c.1843G>A (p.Ala615Thr)	CATCGTCRCCACTGCCCTCTGGG, ACATCGTCRCCACTGCCCTCTGG	Mucopolysaccharidosis, MPS-III-C, RETINITIS PIGMENTOSA 73
121908286	HGSNAT	NM_152419.2(HGSNAT):c.1553C>T (p.Ser518Phe)	CCATTACAGGGGCTCATTTYTGT	Mucopolysaccharidosis, MPS-III-C
397514493	HINT1	NM_005340.6(HINT1):c.278G>A (p.Gly93Asp)	GAATAAGGRTTATCGAATGGTGG	Gamstorp-Wohlfart syndrome
397514492	HINT1	NM_005340.6(HINT1):c.184C>T (p.Gln62Ter)	CCCAAGAAACATATATCCYAGAT, CCAAGAAACATATATCCYAGATT	Gamstorp-Wohlfart syndrome
146448211	HLCS	NM_000411.6(HLCS):c.1993C>T (p.Arg665Ter)	AGTATCRGTAATAAAGGGGAAGG	not provided
119103231	HLCS	NM_000411.6(HLCS):c.1648G>A (p.Val550Met)	TGGCTGTCRTGGAAGCAGTGAGG	Holocarboxylase synthetase deficiency, not provided
119103229	HLCS	NM_000411.6(HLCS):c.1522C>T (p.Arg508Trp)	CCTGTGTTCCAGGAYGGGGAGGG	Holocarboxylase synthetase deficiency
118204096	HMBS	NM_000190.3(HMBS):c.518G>A (p.Arg173Gln)	CAACACCCRGCTTCGGAAGCTGG	Acute intermittent porphyria
118204100	HMBS	NM_000190.3(HMBS):c.593G>A (p.Trp198Ter)	TGGGCTRGCACAACCGGGTGGGG, ATGGGCTRGCACAACCGGGTGGG, CATGGGCTRGCACAACCGGGTGG	Acute intermittent porphyria

118204103	HMBS	NM_000190.3(HMBS):c.77G>A (p.Arg26His)	GGTACCCRCAAGAGCCAGGTGGG, GGGTACCCRCAAGAGCCAGGTGG	Acute intermittent porphyria
118204104	HMBS	NM_000190.3(HMBS):c.91G>A (p.Ala31Thr)	GCAGCTTRCTCGCATACAGACGG	Acute intermittent porphyria
118204110	HMBS	NM_000190.3(HMBS):c.667G>A (p.Glu223Lys)	GGGCGTGRAAGTGCGAGCCAAGG	Acute intermittent porphyria
118204112	HMBS	NM_000190.3(HMBS):c.748G>A (p.Glu250Lys)	TCGCTRAAAGGGCCTTCCTGAGG	Acute intermittent porphyria
118204113	HMBS	NM_000190.3(HMBS):c.754G>A (p.Ala252Thr)	AAGGRCCTTCCTGAGGCACCTGG	Acute intermittent porphyria
118204116	HMBS	NM_000190.3(HMBS):c.647G>A (p.Gly216Asp)	TGGRCCAGGTACACTTGACCAGG	Acute intermittent porphyria
118204094	HMBS	NM_000190.3(HMBS):c.346C>T (p.Arg116Trp)	CCTTCCCTCCTCCCCCAGGYGGG, CCCTCCTCCCCCAGGYGGGAAAA, CCTCCTCCCCCAGGYGGGAAAAC	Acute intermittent porphyria
118204101	HMBS	NM_000190.3(HMBS):c.499C>T (p.Arg167Trp)	CCTTAGCAACTCTCCACAGYGGG	Acute intermittent porphyria
28937320	HMGCS2	NM_001166107.1(HMGCS2):c.160G>A (p.Val54Met)	GGACRTGGGCATCCTGGCCCTGG	mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase deficiency
137852639	HMGCS2	NM_001166107.1(HMGCS2):c.1373G>A (p.Arg458His)	AGCATCRCCGAAAGTATGCCCGG	mitochondrial 3-hydroxy-3-methylglutaryl-CoA synthase deficiency
137853238	HNF1A	NM_000545.6(HNF1A):c.815G>A (p.Arg272His)	GGCRCAAAGAAGAAGCCTTCCGG	Diabetes mellitus, insulin-dependent, 20
137853241	HNF1A	NM_000545.6(HNF1A):c.1859C>T (p.Thr620Ile)	CCACAGCGTCATCGAGAYCTTCA	Maturity-onset diabetes of the young, type 3
137853243	HNF1A	NM_000545.6(HNF1A):c.335C>T (p.Pro112Leu)	CCCTCTCCCAGGGAGGACCYGTG, CCTCTCCCAGGGAGGACCYGTGG	Maturity-onset diabetes of the young, type 3
121918675	HNF1B	NM_000458.3(HNF1B):c.494G>A (p.Arg165His)	AGCRTGCCGCTCTGTACACCTGG	Familial hypoplastic, glomerulocystic kidney
137853336	HNF4A	NM_000457.4(HNF4A):c.406C>T (p.Arg136Trp)	CCAGAATGAGCGGGACYGGATCA	Maturity-onset diabetes of the young, type 1
777046879	HOGA1	NM_138413.3(HOGA1):c.973G>A (p.Gly325Ser)	CAACRGCTGGCTCTGAGGGCAGG, CCAGCAACRGCTGGCTCTGAGGG	Primary hyperoxaluria, type III
764396564	HOGA1	NM_138413.3(HOGA1):c.1	CCCCCCTGTGACCACCCYCTTCA,	Primary hyperoxaluria, type III

		34C>T (p.Pro45Leu)	CCCCCTGTGACCACCCYCTTCAC, CCCCTGTGACCACCCYCTTCACT, CCCTGTGACCACCCYCTTCACTG	
104894019	HOXA13	NM_000522.4(HOXA13):c.1107G>A (p.Trp369Ter)	AATCTGRTTCCAGAACAGGAGGG, CAATCTGRTTCCAGAACAGGAGG	Hand foot uterus syndrome
550921485	HPCA	NM_002143.2(HPCA):c.568G>A (p.Ala190Thr)	CAGCRCCTCCCAGTTCTGAGAGG	Dystonia 2, torsion, autosomal recessive
398123240	HPRT1	NM_000194.2(HPRT1):c.384+1G>A	GAAAGRTATGTATCTTGAAAGGG, GGAAAGRTATGTATCTTGAAAGG	not provided
398123241	HPRT1	NM_000194.2(HPRT1):c.486-1G>A	TTAACARCTTGCTGGTGAAAAGG	not provided
137852506	HPRT1	NM_000194.2(HPRT1):c.193C>T (p.Leu65Phe)	CCATCACATTGTAGCCYTCTGTG	Partial hypoxanthine-guanine phosphoribosyltransferase deficiency
281865089	HPS1	NM_000195.4(HPS1):c.1749G>A (p.Trp583Ter)	CCAGCTGGATCAGAGAYCAGACC	Hermansky-Pudlak syndrome 1
121908316	HPS3	NM_032383.4(HPS3):c.1189C>T (p.Arg397Trp)	CCTGCAGTGTTTCACTGTGYGGT	Hermansky-Pudlak syndrome 3
119471023	HPS4	NM_022081.5(HPS4):c.649C>T (p.Arg217Ter)	CCAAGGTCCTGCTTCACYGAACA	Hermansky-Pudlak syndrome 4
281865107	HPS6	NM_024747.5(HPS6):c.223C>T (p.Gln75Ter)	GGAGGGCTRGCCGGCCGGCCAGG	Hermansky-Pudlak syndrome 6
281865109	HPS6	NM_024747.5(HPS6):c.815C>T (p.Thr272Ile)	GCCCAGRTGTGTACAGCCAGTGG	Hermansky-Pudlak syndrome 6
281865112	HPS6	NM_024747.5(HPS6):c.1234C>T (p.Gln412Ter)	CCGCCGCTRGTAGTACCCGCAGG	Hermansky-Pudlak syndrome 6
121434451	HR	NM_005144.4(HR):c.3034G>A (p.Asp1012Asn)	GGCCRACCTGGTCAGCATCCTGG	Alopecia universalis congenita
121917780	HSD11B2	NM_000196.3(HSD11B2):c.622C>T (p.Arg208Cys)	CCTCCTGCCCCTGCTGYGCAGCT	Apparent mineralocorticoid excess
28935475	HSD17B10	NM_001037811.2(HSD17B10):c.388C>T (p.Arg130Cys)	CCTTCAATGTGATCYGCCTGGTG	2-methyl-3-hydroxybutyric aciduria
119481078	HSD17B3	NM_000197.1(HSD17B3):c.166G>A (p.Ala56Thr)	CTGGARCAGGCGATGGAATTGGG, ACTGGARCAGGCGATGGAATTGG	Testosterone 17-beta-dehydrogenase deficiency
28939085	HSD17B3	NM_000197.1(HSD17B3):c.695C>T (p.Ser232Leu)	CCCCATATGCTGTCTYGACTGCA, CCCATATGCTGTCTYGACTGCAA	Testosterone 17-beta-dehydrogenase deficiency

80358216	HSD3B2	NM_001166120.1(HSD3B2):c.512G>A (p.Trp171Ter)	AATGGGTRGAATCTAAAAAATGG	3 beta-Hydroxysteroid dehydrogenase deficiency
28937569	HSPB1	NM_001540.3(HSPB1):c.545C>T (p.Pro182Leu)	CCAACGAGATCACCATCCYAGTC	Distal hereditary motor neuronopathy type 2B
104894020	HSPB1	NM_001540.3(HSPB1):c.544C>T (p.Pro182Ser)	CCAACGAGATCACCATCYCAGTC	Distal hereditary motor neuronopathy type 2B
29001571	HSPB1	NM_001540.3(HSPB1):c.379C>T (p.Arg127Trp)	CCGGCAAGCACGAGGAGYGGCAG	Charcot-Marie-Tooth disease type 2F, Distal hereditary motor neuronopathy type 2B
137853248	HSPG2	NM_005529.6(HSPG2):c.4595G>A (p.Cys1532Tyr)	CGCTRCCCGCCAGGCTACATCGG	Schwartz Jampel syndrome type 1
587776445	HTRA1	NM_002775.4(HTRA1):c.821G>A (p.Arg274Gln)	GCTGCRGCCGGGAGAGTTCGTGG	Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy
113993969	HTRA1	NM_002775.4(HTRA1):c.889G>A (p.Val297Met)	ATCRTGAGCACCACCCAGCGAGG	Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy
113993971	HTRA1	NM_002775.4(HTRA1):c.1108C>T (p.Arg370Ter)	CCTCACGGAGTCCCATGACYGAC	Cerebral autosomal recessive arteriopathy with subcortical infarcts and leukoencephalopathy
104893743	HYAL1	NM_153281.1(HYAL1):c.802G>A (p.Glu268Lys)	TGTGGCCRAGGCATTCCGTGTGG	Deficiency of hyaluronoglucosaminidase
373436822	IARS2	NM_018060.3(IARS2):c.1821G>A (p.Trp607Ter)	TCATGRTCTTATGTTCTTCCAGG	Leigh disease, not provided
118203918	ICK	NM_016513.4(ICK):c.815G>A (p.Arg272Gln)	GAAACRACCAACAGCTAGTCAGG	Endocrine-cerebroosteodysplasia
121913500	IDH1	NM_001282386.1(IDH1):c.395G>A (p.Arg132His)	CATAGGTCRTCATGCTTATGGGG	
121913502	IDH2	NM_002168.3(IDH2):c.419G>A (p.Arg140Gln)	CTATCCRGAACATCCTGGGGGGG, ACTATCCRGAACATCCTGGGGGG, AACTATCCRGAACATCCTGGGGG	D-2-hydroxyglutaric aciduria 2
104894853	IDS	NM_000202.6(IDS):c.998C>T (p.Ser333Leu)	CCATCATTGCATTTACCTYGGAT	Mucopolysaccharidosis, MPS-II, not provided
199422231	IDS	NM_000202.6(IDS):c.1402C>T (p.Arg468Trp)	CCTATAGCCAGTATCCCYGGCCT	Mucopolysaccharidosis, MPS-II
121965019	IDUA	NM_000203.4(IDUA):c.1205G>A (p.Trp402Ter)	GCTCTRGGCCGAAGTGTCGCAGG	Hurler syndrome, not provided
121965030	IDUA	NM_000203.4(IDUA):c.898G>A (p.Ala300Thr)	GACGAGRCGGACCCGCTGGTGGG, CGACGAGRCGGACCCGCTGGTGG	
121965032	IDUA	NM_000203.4(IDUA):c.109	CCCCTTCGCGCAGCGCAYGCTCA,	Mucopolysaccharidosis, MPS-I-H/S

		1C>T (p.Thr364Met)	CCCTTCGCGCAGCGCAYGCTCAC, CCTTCGCGCAGCGCAYGCTCACC	
786201032	IFITM5	NM_001025295.2(IFITM5) :c.119C>T (p.Ser40Leu)	CCACTTGATCTGGTYGGTGTTCA	Osteogenesis imperfecta type 5
431905521	IFT140	NM_014714.3(IFT140):c.874G>A (p.Val292Met)	TCTCRTGATGGCCGTCGGGGAGG, CCTTCTCRTGATGGCCGTCGGGG, GCCTTCTCRTGATGGCCGTCGGG	Renal dysplasia, retinal pigmentary dystrophy, cerebellar ataxia and skeletal dysplasia
794727473	IFT140	NM_014714.3(IFT140):c.3991C>T (p.Gln1331Ter)	CCAGGAGACCAGGCTGGCGYAGC	Renal dysplasia, retinal pigmentary dystrophy, cerebellar ataxia and skeletal dysplasia
199826737	IFT140	NM_014714.3(IFT140):c.1565G>A (p.Gly522Glu)	CCAAGAAGCAGGGATTCYCCTCA	
201188361	IFT140	NM_014714.3(IFT140):c.634G>A (p.Gly212Arg)	CCAAAGTTCCTCACYGTCCATCA	Renal dysplasia, retinal pigmentary dystrophy, cerebellar ataxia and skeletal dysplasia
145541911	IFT172	NM_015662.2(IFT172):c.886C>T (p.Arg296Trp)	GCCATCCCRCTTCCAGGCCAAGG	
587777546	IFT27	NM_001177701.2(IFT27):c.299G>A (p.Cys100Tyr)	CCAGCCACTTGCTGYAGTTGTTG	Bardet-Biedl syndrome 19
137852667	IGHMBP2	NM_002180.2(IGHMBP2): c.1738G>A (p.Val580Ile)	TTCRTCAGATCCAACAGGAAAGG, TGTCCTTCRTCAGATCCAACAGG	Werdnig-Hoffmann disease, Charcot-Marie-Tooth disease
74315491	IGLL1	NM_020070.3(IGLL1):c.64C>T (p.Gln22Ter)	CCAGGCCCCAACCTCAGGYAGCG	Agammaglobulinemia 2, autosomal recessive
121917853	IHH	NM_002181.3(IHH):c.391G>A (p.Glu131Lys)	ACCRAGGGCTGGGACGAGGACGG, GGTGACCRAGGGCTGGGACGAGG	Brachydactyly type A1
121917855	IHH	NM_002181.3(IHH):c.298G>A (p.Asp100Asn)	CGCCRACCGCCTCATGACCCAGG	Brachydactyly type A1
267606873	IHH	NM_002181.3(IHH):c.383G>A (p.Arg128Gln)	GCTGCRGGTGACCGAGGGCTGGG, AGCTGCRGGTGACCGAGGGCTGG	Brachydactyly type A1
121917861	IHH	NM_002181.3(IHH):c.461C>T (p.Thr154Ile)	CCGCGCGGTGGACATCAYCACAT	Brachydactyly type A1
121917856	IHH	NM_002181.3(IHH):c.137C>T (p.Pro46Leu)	CCGCCACGCAAACTCGTGCYGCT, CCACGCAAACTCGTGCYGCTCGC	Acrocapitofemoral dysplasia
200296680	IKBKB	NM_001556.2(IKBKB):c.814C>T (p.Arg272Ter)	CCACAGTGTCCTGGCTGAGYGAC	Immunodeficiency 15
137853329	IKBKG	NM_003639.4(IKBKG):c.1207C>T (p.Gln403Ter)	CCCAAGTGCCAGTATYAGGCCCC, CCAAGTGCCAGTATYAGGCCCCT	Hypohidrotic ectodermal dysplasia with immune deficiency

149491038	IL10RA	NM_001558.3(IL10RA):c.784C>T (p.Arg262Cys)	CCTCCAGCTGTATGTGCGGYGCC, CCAGCTGTATGTGCGGYGCCGAA	
137853580	IL10RA	NM_001558.3(IL10RA):c.251C>T (p.Thr84Ile)	CCCTGTCCTATGACCTTAYCGCA, CCTGTCCTATGACCTTAYCGCAG	
387906787	IL11RA	NM_001142784.2(IL11RA):c.475C>T (p.Gln159Ter)	CCTAGGAGCTGATAGCYAGAGGT	Craniosynostosis and dental anomalies
121434492	IL12RB1	NM_005535.2(IL12RB1):c.94C>T (p.Gln32Ter)	CCAGTGAGTGCTGTTTTYAGGAC	Immunodeficiency 30
748486078	IL17F	NM_052872.3(IL17F):c.284C>T (p.Ser95Leu)	CTTCCRAGGGGTACCGGTTGGGG, ACTTCCRAGGGGTACCGGTTGGG, AACTTCCRAGGGGTACCGGTTGG	Candidiasis, familial, 6
387906913	IL17RA	NM_014339.6(IL17RA):c.850C>T (p.Gln284Ter)	CCCTCTCTGCCCGCAGATCYAGC, CCTCTCTGCCCGCAGATCYAGCC	Candidiasis, familial, 5
122461161	IL1RAPL1	NM_014271.3(IL1RAPL1):c.1460G>A (p.Trp487Ter)	GGGCTRGAGCATCTTTGAGCTGG	Mental retardation 21, X-linked
137852508	IL2RG	NM_000206.2(IL2RG):c.865C>T (p.Arg289Ter)	CCCTGTCAGGACGATGCCCYGAA, CCTGTCAGGACGATGCCCYGAAT	X-linked severe combined immunodeficiency
104893894	IL7R	NM_002185.3(IL7R):c.394C>T (p.Pro132Ser)	CCAGTTAAACCTGAGGCTYCTTT	Severe combined immunodeficiency, autosomal recessive, T cell-negative, B cell-positive, NK cell-positive
121912551	IMPDH1	NM_000883.3(IMPDH1):c.1057G>A (p.Val353Ile)	GGGCRTCGACGTCATAGTCTTGG	Retinitis pigmentosa 10
121912553	IMPDH1	NM_000883.3(IMPDH1):c.568C>T (p.Arg190Trp)	CCAGGCCAACGAGGTGYGGAAGG	Leber congenital amaurosis 11
267607183	INF2	NM_022489.3(INF2):c.653G>A (p.Arg218Gln)	CAGCTGCRGAACGAGTTTATCGG	Focal segmental glomerulosclerosis 5
121918129	INPP5E	NM_019892.4(INPP5E):c.1304G>A (p.Arg435Gln)	AGCRGCTGCTGGACTACACCAGG	Familial aplasia of the vermis, Joubert syndrome 1
104894698	INSL3	NM_005543.3(INSL3):c.304C>T (p.Arg102Cys)	CCAGACCTCTCACCATCACYGCC, CCTCTCACCATCACYGCCACCAC	Cryptorchidism, unilateral or bilateral
121913139	INSR	NM_000208.2(INSR):c.3481G>A (p.Ala1161Thr)	CCTGRCAGCGAGAAACTGCATGG	Insulin resistance, Insulin-resistant diabetes mellitus AND acanthosis nigricans
121913146	INSR	NM_000208.2(INSR):c.479G>A (p.Trp160Ter)	TATCGACTRGTCCCGTATCCTGG	Insulin-resistant diabetes mellitus AND acanthosis nigricans
121913150	INSR	NM_000208.2(INSR):c.3572G>A (p.Arg1191Gln)	TACTACCRGAAAGGGGGCAAGGG, TTACTACCRGAAAGGGGGCAAGG	Diabetes mellitus type 2
121913156	INSR	NM_000208.2(INSR):c.3602G>A (p.Arg1201Gln)	CCCTGTACRGTGGATGGCACCGG	Insulin-resistant diabetes mellitus AND acanthosis nigricans, Hyperinsulinemic hypoglycemia familial 5

1799816	INSR	NM_000208.2(INSR):c.3034G>A (p.Val1012Met)	GTACRTGCCGGACGAGTGGGAGG, TGTGTACRTGCCGGACGAGTGGG, CTGTGTACRTGCCGGACGAGTGG	Diabetes mellitus type 2, not specified
755549444	INVS	NM_014425.3(INVS):c.2509C>T (p.Gln837Ter)	CCAAGAAACAAAGTGACAYAAGC	Infantile nephronophthisis
121918244	IQCB1	NM_001023570.2(IQCB1):c.1381C>T (p.Arg461Ter)	CCGAGTTGAACTGAAGAAAYGAG	Senior-Loken syndrome 5, not provided
727503968	IQCB1	NM_001023570.2(IQCB1):c.1090C>T (p.Arg364Ter)	CCATGAGACTTTCCYGAGAATTG	Senior-Loken syndrome 5, not provided
587777261	IQSEC2	NM_001111125.2(IQSEC2):c.2563C>T (p.Arg855Ter)	GATGAGTCRCTCCACTTTCTGGG	Mental retardation, X-linked, nonspecific
267607186	IQSEC2	NM_001111125.2(IQSEC2):c.2587C>T (p.Arg863Trp)	CCTCCTGCCCTGCAGCCAGYGGT, CCTGCCCTGCAGCCAGYGGTACT	Mental retardation, X-linked, nonspecific, not provided
267607188	IQSEC2	NM_001111125.2(IQSEC2):c.1075C>T (p.Arg359Cys)	CCATCCAGACAGCCTTCYGCCAG	Mental retardation, X-linked, nonspecific
377584435	IRAK4	NM_016123.3(IRAK4):c.34C>T (p.Arg12Cys)	CCATCAACATATGTGYGCTGCCT	IRAK4 deficiency
121434228	IRF6	NM_006147.3(IRF6):c.1137G>A (p.Trp379Ter)	ATGRCCAGATGGGAAACCATTGG	Van der Woude syndrome
397515434	IRF6	NM_006147.3(IRF6):c.145C>T (p.Gln49Ter)	CCACCCGGCATAGCCCTYAACAA, CCCGGCATAGCCCTYAACAAGAA	Van der Woude syndrome
28942093	IRF6	NM_006147.3(IRF6):c.5C>T (p.Ala2Val)	CCCCCCCAGATCATGGYCCTCCA, CCCCCCAGATCATGGYCCTCCAC, CCCCCAGATCATGGYCCTCCACC	Van der Woude syndrome
121434230	IRF6	NM_006147.3(IRF6):c.1186C>T (p.Pro396Ser)	CCTGAACAGGTCATTYCAGTAGT	Van der Woude syndrome
387906968	IRF6	NM_006147.3(IRF6):c.1271C>T (p.Ser424Leu)	CCGCCTGCAGATCTYAACCCCAG	Popliteal pterygium syndrome
786201005	ISG15	NM_005101.3(ISG15):c.163C>T (p.Gln55Ter)	CCCGAGCGGTGTGGCGCTGYAGG, CCGAGCGGTGTGGCGCTGYAGGA	Immunodeficiency 38
368593151	ISPD	NM_001101426.3(ISPD):c.802C>T (p.Arg268Ter)	CTCRTTTGTAGGTCACCTAAAGG	Congenital muscular dystrophy-dystroglycanopathy with brain and eye anomalies, type A7
137852907	ITGA2B	NM_000419.3(ITGA2B):c.818G>A (p.Gly273Asp)	GGCCGTGGRCGAGTTCGACGGGG	Glanzmann thrombasthenia
137852906	ITGA2B	NM_000419.3(ITGA2B):c.	CCACCATGGCCTTCCTTYGAGTA,	Glanzmann thrombasthenia

		1750C>T (p.Arg584Ter)	CCATGGCCTTCCTTYGAGTACGC	
200402328	ITGA7	NM_002206.2(ITGA7):c.2357+1G>A	CCGGCCCCGCCTGGCTTAYGTGG, CCCCGCCTGGCTTAYGTGGCCAA	Congenital muscular dystrophy
374664941	ITGA8	NM_003638.2(ITGA8):c.1219G>A (p.Gly407Arg)	CCTGCAAAAGGCACTCYGATGGC	Renal adysplasia
9983887	ITGB2	NM_000211.4(ITGB2):c.329-6C>T	GCCRGTGGGGACAGAACAAAAGG	Leukocyte adhesion deficiency type 1
137852616	ITGB2	NM_000211.4(ITGB2):c.850G>A (p.Gly284Ser)	CGACRGCCGCTGTCACCTGGAGG, CAACGACRGCCGCTGTCACCTGG	Leukocyte adhesion deficiency
121918449	ITGB3	NM_000212.2(ITGB3):c.1199G>A (p.Cys400Tyr)	TGCCACCTRCCTCAACAATGAGG	Glanzmann thrombasthenia
121912466	ITGB4	NM_000213.3(ITGB4):c.2792G>A (p.Gly931Asp)	CCGGGRCATGGTGGAGTTCCAGG	Adult junctional epidermolysis bullosa
80338755	ITGB4	NM_000213.3(ITGB4):c.182G>A (p.Cys61Tyr)	CCGGCGCTRCAACACCCAGGCGG	Epidermolysis bullosa with pyloric atresia
121912462	ITGB4	NM_000213.3(ITGB4):c.1660C>T (p.Arg554Ter)	CCCTCTCTGCAGACYGAGGACGC	Epidermolysis bullosa with pyloric atresia
121908191	ITK	NM_005546.3(ITK):c.1003C>T (p.Arg335Trp)	CCTGGTGACTCGACTCYGGTATC	Lymphoproliferative syndrome 1
763471771	IVD	NM_002225.3(IVD):c.793+1G>A	GATTCCTGRTAAGTAGCACCGGG	Isovaleryl-CoA dehydrogenase deficiency
28940889	IVD	NM_002225.3(IVD):c.941C>T (p.Ala314Val)	CCTGCACGTGAGGGAAGYCTTTG	Isovaleryl-CoA dehydrogenase deficiency, not provided
28939668	JAG1	NM_000214.2(JAG1):c.821G>A (p.Gly274Asp)	ACGRCATCTGTAATGAGCCCTGG	Tetralogy of Fallot
121918350	JAG1	NM_000214.2(JAG1):c.550C>T (p.Arg184Cys)	CCACTTTGAGTATCAGATCYGCG	Alagille syndrome 1
587777727	JAGN1	NM_032492.3(JAGN1):c.3G>A (p.Met1Ile)	GGCACAATRGCGTCTCGAGCAGG	Severe congenital neutropenia, Severe congenital neutropenia 6, autosomal recessive
137852626	JAK3	NM_000215.3(JAK3):c.1333C>T (p.Arg445Ter)	CCTTCTGGTTGGCCTCAGCYGAC	Severe combined immunodeficiency, autosomal recessive, T cell-negative, B cell-positive, NK cell-negative
796052595	KANSL1	NM_001193466.1(KANSL1):c.2203+1G>A	CCARTAAGTGTCAGGGAGCCGGG, GCCARTAAGTGTCAGGGAGCCGG	not provided
397514746	KARS	NM_001130089.1(KARS):c.1129G>A (p.Asp377Asn)	TCACRATCTCATGGAAATCACGG	Deafness, autosomal recessive 89
730880257	KATNB1	NM_005886.2(KATNB1):c.1604C>T (p.Ser535Leu)	CCATCAACGACCTGTYGGTGGTG	Lissencephaly 6, with microcephaly

786205232	KCNA2	NM_004974.3(KCNA2):c.890G>A (p.Arg297Gln)	TCCRGTTGGTAAGAGTCTTTAGG	EPILEPTIC ENCEPHALOPATHY, EARLY INFANTILE, 32
121908593	KCNA5	NM_002234.3(KCNA5):c.1828G>A (p.Glu610Lys)	CCGGRAAACAGATTTGTGAAAGG	Atrial fibrillation, familial, 7
1805128	KCNE1	NM_000219.5(KCNE1):c.253G>A (p.Asp85Asn)	GTCCRATGCCTGGCAAGAGAAGG	Long QT syndrome, Long QT syndrome 5, acquired, susceptibility to, Long QT syndrome 2/5, not specified, not provided
199473360	KCNE1	NM_000219.5(KCNE1):c.247G>A (p.Glu83Lys)	TCTACATCRAGTCCGATGCCTGG	Congenital long QT syndrome
199473644	KCNE1	NM_000219.5(KCNE1):c.163G>A (p.Gly55Ser)	TTCTTCRGCTTCTTCACCCTGGG, ATTCTTCRGCTTCTTCACCCTGG	Congenital long QT syndrome, not specified
79654911	KCNE1	NM_000219.5(KCNE1):c.200G>A (p.Arg67His)	CCAGCTTCTTGGAGYGGATGTAG	Long QT syndrome, Congenital long QT syndrome
74315446	KCNE1	NM_000219.5(KCNE1):c.221C>T (p.Ser74Leu)	CCAAGAAGCTGGAGCACTYGAAC	Long QT syndrome 5, Congenital long QT syndrome
28933384	KCNE1	NM_000219.5(KCNE1):c.20C>T (p.Thr7Ile)	CCTGTCTAACACCAYAGCGGTGA	Jervell and Lange-Nielsen syndrome 2, Congenital long QT syndrome
199473367	KCNE2	NM_172201.1(KCNE2):c.347C>T (p.Ala116Val)	CCATGAGAACATTGGTGYGGCTG	Acquired long QT syndrome
199473648	KCNE2	NM_172201.1(KCNE2):c.29C>T (p.Thr10Met)	CCAATTTCACACAGAYGCTGGAA	Long QT syndrome, Congenital long QT syndrome, Cardiac arrhythmia
730882174	KCNH1	NM_002238.3(KCNH1):c.1042G>A (p.Gly348Arg)	TGAATATRGAGCTGCTGTGCTGG	Zimmermann-Laband syndrome
794728397	KCNH2	NM_000238.3(KCNH2):c.2770G>A (p.Gly924Arg)	CCGGCCGRGGGGGCCGTGGGGGG, GCCGGCCGRGGGGGCCGTGGGGG	Cardiac arrhythmia
794728401	KCNH2	NM_000238.3(KCNH2):c.3002G>A (p.Trp1001Ter)	AGCTTCTRGGGGGACAGTCGGGG, CAGCTTCTRGGGGGACAGTCGGG	Cardiac arrhythmia
794728478	KCNH2	NM_000238.3(KCNH2):c.1129-1G>A	CGGGTGCARGTCCTGTCCCTGGG	Cardiac arrhythmia
794728487	KCNH2	NM_000238.3(KCNH2):c.1945+1G>A	GCTRTGAGTGTGCCCAGGGGCGG, TTGGCTRTGAGTGTGCCCAGGGG, ATTGGCTRTGAGTGTGCCCAGGG, CATTGGCTRTGAGTGTGCCCAGG	Cardiac arrhythmia
141401803	KCNH2	NM_000238.3(KCNH2):c.2860C>T (p.Arg954Cys)	GGCRGAGGGGGCTGGAGCTGCGG	Sudden infant death syndrome, Cardiac arrhythmia

199472880	KCNH2	NM_000238.3(KCNH2):c.865G>A (p.Glu289Lys)	CATCRAGGCCATGCGCGCCGGGG, ACATCRAGGCCATGCGCGCCGGG, GACATCRAGGCCATGCGCGCCGG	Congenital long QT syndrome
199472937	KCNH2	NM_000238.3(KCNH2):c.1811G>A (p.Gly604Asp)	CCTGGGCGDCCCCTCCATCAAGG	Congenital long QT syndrome
199473019	KCNH2	NM_000238.3(KCNH2):c.3014G>A (p.Arg1005Gln)	CAGTCRGGGCCGCCAGTACCAGG	Congenital long QT syndrome
199473022	KCNH2	NM_000238.3(KCNH2):c.3107G>A (p.Gly1036Asp)	CCCGGGRCGACGTGGAGAGCAGG	Congenital long QT syndrome, Cardiac arrhythmia
199473432	KCNH2	NM_000238.3(KCNH2):c.2660G>A (p.Arg887His)	AGCRCAAGTTGTCCTTCCGCAGG	Long QT syndrome, Congenital long QT syndrome
199473540	KCNH2	NM_000238.3(KCNH2):c.2810G>A (p.Ser937Asn)	CTCCARCCCTGAGAGCAGTGAGG	Congenital long QT syndrome, Cardiac arrhythmia
199473669	KCNH2	NM_000238.3(KCNH2):c.2707G>A (p.Gly903Arg)	CCARGGGAGGTGTCGGCCTTGGG, GCCARGGGAGGTGTCGGCCTTGG	Congenital long QT syndrome, Cardiac arrhythmia
199473670	KCNH2	NM_000238.3(KCNH2):c.2759G>A (p.Arg920Gln)	GTAGCCRGGGCCGGCCGGGGGG, AGTAGCCRGGGCCGGCCGGGGG, GAGTAGCCRGGGCCGGCCGGGG	Congenital long QT syndrome
121912509	KCNH2	NM_000238.3(KCNH2):c.3003G>A (p.Trp1001Ter)	AGCTTCTGRGGGGACAGTCGGGG	Long QT syndrome 2, Cardiac arrhythmia
138498207	KCNH2	NM_000238.3(KCNH2):c.2371C>T (p.Arg791Trp)	GTCGCCCCRCAGGATCTCGATGG	Long QT syndrome, Congenital long QT syndrome, Cardiac arrhythmia
770047651	KCNH2	NM_000238.3(KCNH2):c.1128G>A (p.Gln376=)	CCGGCCGCTGGGCGCCTACYTGG, CCGCTGGGCGCCTACYTGGGTGA	Long QT syndrome, Cardiac arrhythmia
794728381	KCNH2	NM_000238.3(KCNH2):c.2026C>T (p.Gln676Ter)	CCCGCTACCACACAYAGATGCTG	Cardiac arrhythmia
794728403	KCNH2	NM_000238.3(KCNH2):c.3040C>T (p.Arg1014Ter)	CCAGTACCAGGAGCTCCCTYGAT	Cardiac arrhythmia
794728364	KCNH2	NM_000238.3(KCNH2):c.1096C>T (p.Arg366Ter)	CCTAAGATAAAGGAGYGAACCCA	Cardiac arrhythmia
794728481	KCNH2	NM_000238.3(KCNH2):c.1684C>T (p.His562Tyr)	CCTTTGCGCTCATCGCGYACTGG	Cardiac arrhythmia

199472885	KCNH2	NM_000238.3(KCNH2):c.934C>T (p.Arg312Cys)	CCATGCACCCACTGYGCAGCGGC	Long QT syndrome, Congenital long QT syndrome, Cardiac arrhythmia
199472901	KCNH2	NM_000238.3(KCNH2):c.1307C>T (p.Thr436Met)	CCTTCCTGCTGAAGGAGAYGGAA, CCTGCTGAAGGAGAYGGAAGAAG	Congenital long QT syndrome
199472910	KCNH2	NM_000238.3(KCNH2):c.1474C>T (p.His492Tyr)	CCCCGGCCGCATCGCCGTCYACT, CCCGGCCGCATCGCCGTCYACTA, CCGGCCGCATCGCCGTCYACTAC	Congenital long QT syndrome
199472984	KCNH2	NM_000238.3(KCNH2):c.2086C>T (p.Arg696Cys)	CCCCAATCCCCTGCGCCAGYGCC, CCCAATCCCCTGCGCCAGYGCCT, CCAATCCCCTGCGCCAGYGCCTC	Congenital long QT syndrome
199473021	KCNH2	NM_000238.3(KCNH2):c.3097C>T (p.Arg1033Trp)	CCTCTCCAGCCCGGGTCGGYGGC, CCAGCCCGGGTCGGYGGCCCCGG	Congenital long QT syndrome
199473035	KCNH2	NM_000238.3(KCNH2):c.3457C>T (p.His1153Tyr)	CCCAGCCCCTGCACAGAYACGGC, CCAGCCCCTGCACAGAYACGGCT	Congenital long QT syndrome
121912504	KCNH2	NM_000238.3(KCNH2):c.1682C>T (p.Ala561Val)	CCTTTGCGCTCATCGYGCACTGG	Long QT syndrome 2, Congenital long QT syndrome, Cardiac arrhythmia
121912508	KCNH2	NM_000238.3(KCNH2):c.1744C>T (p.Arg582Cys)	CCACACATGGACTCAYGCATCGG	Long QT syndrome 2, Congenital long QT syndrome, Cardiac arrhythmia
794728382	KCNH2	NM_000238.3(KCNH2):c.2104C>T (p.Gln702Ter)	CCTCGAGGAGTACTTCYAGCACG	Cardiac arrhythmia
150988911	KCNH2	NM_000238.3(KCNH2):c.343G>A (p.Val115Met)	CCCCATCCTCGTTCTTCAYGGGC, CCCATCCTCGTTCTTCAYGGGCA, CCATCCTCGTTCTTCAYGGGCAC	Long QT syndrome 2, Congenital long QT syndrome
77331749	KCNH2	NM_000238.3(KCNH2):c.2738C>T (p.Ala913Val)	CCTTGGGGCCGGGCCGGGYGGGG	Long QT syndrome 2, Congenital long QT syndrome, Long QT syndrome 2/9, digenic, Cardiac arrhythmia
104894253	KCNJ1	NM_153767.3(KCNJ1):c.535G>A (p.Ala179Thr)	AACRCAGTGATCAGCAAACGGG, GAACRCAGTGATCAGCAAACGG, AGAACRCAGTGATCAGCAAACGG	Bartter syndrome antenatal type 2
137853067	KCNJ10	NM_002241.4(KCNJ10):c.595C>T (p.Arg199Ter)	CCCTGCCTCATGATCYGAGTTGC, CCTGCCTCATGATCYGAGTTGCC	SeSAME syndrome
137853071	KCNJ10	NM_002241.4(KCNJ10):c.889C>T (p.Arg297Cys)	CCACCTGTCAGGTGYGCACTTCC	SeSAME syndrome
137853074	KCNJ10	NM_002241.4(KCNJ10):c.1042C>T (p.Arg348Cys)	CCTCCGTGACAGCACTGTAYGCT, CCGTGACAGCACTGTAYGCTACG	Enlarged vestibular aqueduct syndrome

80356615	KCNJ11	NM_000525.3(KCNJ11):c.158G>A (p.Gly53Asp)	GGAGCAGGRCCGCTTCCTGCAGG	Permanent neonatal diabetes mellitus, Diabetes mellitus, permanent neonatal, with neurologic features
80356616	KCNJ11	NM_000525.3(KCNJ11):c.175G>A (p.Val59Met)	GGACRTGTTCACCACGCTGGTGG, GCAGGACRTGTTCACCACGCTGG	Permanent neonatal diabetes mellitus, Diabetes mellitus, permanent neonatal, with neurologic features, Neonatal insulin-dependent diabetes mellitus
267607196	KCNJ11	NM_000525.3(KCNJ11):c.844G>A (p.Glu282Lys)	CCTCRAGATCATCGTCATCCTGG	Islet cell hyperplasia
80356625	KCNJ11	NM_000525.3(KCNJ11):c.601C>T (p.Arg201Cys)	CCGCCTCTGCTTCATGCTAYGTG, CCTCTGCTTCATGCTAYGTGTGG	Permanent neonatal diabetes mellitus, Diabetes mellitus, permanent neonatal, with neurologic features, Diabetes mellitus
527236152	KCNJ18	NM_001194958.2(KCNJ18):c.419C>T (p.Thr140Met)	CCTCTTCTCCATCGAGAYGCAGA	Thyrotoxic periodic paralysis
527236157	KCNJ18	NM_001194958.2(KCNJ18):c.1219C>T (p.Gln407Ter)	CCGGGATGGCCTCAGCCCCYAGG	Thyrotoxic periodic paralysis
527236158	KCNJ18	NM_001194958.2(KCNJ18):c.1061C>T (p.Thr354Met)	CCTATGAGGTGCCCTCTAYGCCC	Thyrotoxic periodic paralysis, Thyrotoxic periodic paralysis 2
104894584	KCNJ2	NM_000891.2(KCNJ2):c.514G>A (p.Asp172Asn)	ATCATCRATGCTTTCATCATTGG	Short QT syndrome 3, short QT syndrome
104894581	KCNJ2	NM_000891.2(KCNJ2):c.557C>T (p.Pro186Leu)	CCAAGATGGCAAAGCYAAAGAAG	Andersen Tawil syndrome, Congenital long QT syndrome
786204795	KCNJ6	NM_002240.4(KCNJ6):c.460G>A (p.Gly154Ser)	CACCATTRGTTATGGCTACCGGG, CCACCATTRGTTATGGCTACCGG	Keppen-Lubinsky syndrome
121908332	KCNK9	NM_001282534.1(KCNK9):c.706G>A (p.Gly236Arg)	CATCRGGGCCTTCCTCAACCTGG	Birk Barel mental retardation dysmorphism syndrome
794728531	KCNQ1	NM_000218.2(KCNQ1):c.1685+1G>A	AGGAGRTGGGCACGGCCAAACGG	Cardiac arrhythmia
794728539	KCNQ1	NM_000218.2(KCNQ1):c.1794G>A (p.Lys598=)	GACAARGTAGGCTCACGCGCCGG	Cardiac arrhythmia
794728553	KCNQ1	NM_000218.2(KCNQ1):c.343G>A (p.Glu115Lys)	TCCTCRAGCGTCCCACCGGCTGG	Cardiac arrhythmia
794728572	KCNQ1	NM_000218.2(KCNQ1):c.1176G>A (p.Trp392Ter)	CTGRAAGATCTACATCCGGAAGG, CCACCTGRAAGATCTACATCCGG	Cardiac arrhythmia
17215479	KCNQ1	NM_000218.2(KCNQ1):c.643G>A (p.Val215Met)	CTGCRTGGGCTCCAAGGGGCAGG, TCCTCTGCRTGGGCTCCAAGGGG	Congenital long QT syndrome, Long QT syndrome, LQT1 subtype
120074183	KCNQ1	NM_000218.2(KCNQ1):c.1034G>A (p.Gly345Glu)	CCCAGGRGATTCTTGGCTCGGGG, TCCCAGGRGATTCTTGGCTCGGG, TTCCCAGGRGATTCTTGGCTCGG	Long QT syndrome 1, Congenital long QT syndrome, Long QT syndrome, LQT1 subtype

120074187	KCNQ1	NM_000218.2(KCNQ1):c.898G>A (p.Ala300Thr)	TACRCAGATGCGCTGTGGTGGGG, CTACRCAGATGCGCTGTGGTGGG, GCTACRCAGATGCGCTGTGGTGG, GCAGCTACRCAGATGCGCTGTGG	Long QT syndrome 1, Cardiac arrhythmia, not provided
120074188	KCNQ1	NM_000218.2(KCNQ1):c.1573G>A (p.Ala525Thr)	TGTGRCCAAGAAGAAATTCCAGG	Congenital long QT syndrome, Cardiac arrhythmia, Long QT syndrome, LQT1 subtype
199472688	KCNQ1	NM_000218.2(KCNQ1):c.436G>A (p.Glu146Lys)	CATCRAGCAGTATGCCGCCCTGG	Congenital long QT syndrome, Long QT syndrome, LQT1 subtype
199472692	KCNQ1	NM_000218.2(KCNQ1):c.484G>A (p.Val162Met)	ATCRTGCTGGTGGTGTTCTTCGG	Congenital long QT syndrome, Cardiac arrhythmia
199472694	KCNQ1	NM_000218.2(KCNQ1):c.514G>A (p.Val172Met)	CGGAGTACRTGGTCCGCCTCTGG	Congenital long QT syndrome
199472736	KCNQ1	NM_000218.2(KCNQ1):c.875G>A (p.Gly292Asp)	GAGTCAGRCCGCGTGGAGTTCGG	Congenital long QT syndrome, Cardiac arrhythmia
199472811	KCNQ1	NM_000218.2(KCNQ1):c.1750G>A (p.Gly584Ser)	GATCGCRGCAGCAACACGATCGG	Sudden infant death syndrome
199473464	KCNQ1	NM_000218.2(KCNQ1):c.868G>A (p.Glu290Lys)	GGTGAACRAGTCAGGCCGCGTGG	Congenital long QT syndrome, Long QT syndrome, LQT1 subtype
199473482	KCNQ1	NM_000218.2(KCNQ1):c.1748G>A (p.Arg583His)	GATCRCGGCAGCAACACGATCGG	Congenital long QT syndrome, Cardiac arrhythmia
762814879	KCNQ1	NM_000218.2(KCNQ1):c.477+1G>A	GGATGRTACGTAGCATCTGAGGG, TGGATGRTACGTAGCATCTGAGG	Cardiac arrhythmia
151344631	KCNQ1	NM_000218.2(KCNQ1):c.613G>A (p.Val205Met)	CATCRTGGTCGTGGCCTCCATGG	Long QT syndrome 1, Long QT syndrome, Congenital long QT syndrome, Cardiac arrhythmia, Long QT syndrome, LQT1 subtype, not provided
149089817	KCNQ1	NM_000218.2(KCNQ1):c.1336G>A (p.Asp446Asn)	CGTGCRACCCCCCAGAAGAGCGG	Cardiac arrhythmia
397508070	KCNQ1	NM_000218.2(KCNQ1):c.1032+1G>A	CCCAGCGRTAGGTGCCCCGTGGG, TCCCAGCGRTAGGTGCCCCGTGG	Long QT syndrome, Cardiac arrhythmia, Long QT syndrome, LQT1 subtype
140452381	KCNQ1	NM_000218.2(KCNQ1):c.1354C>T (p.Arg452Trp)	CCCCCCAGAAGAGCGGYGGCTGG, CCCCCAGAAGAGCGGYGGCTGGA, CCCCAGAAGAGCGGYGGCTGGAC	Long QT syndrome, Congenital long QT syndrome, Cardiac arrhythmia, Long QT syndrome, LQT1 subtype
199472787	KCNQ1	NM_000218.2(KCNQ1):c.1555C>T (p.Arg519Cys)	CCATTAAGGTCATTCGAYGCATG	Congenital long QT syndrome

199473446	KCNQ1	NM_000218.2(KCNQ1):c.197C>T (p.Ser66Phe)	CCCGCGCCCCCTGCGTYCCCGGC, CCGCGCCCCCTGCGTYCCCGGCC	Congenital long QT syndrome
199473450	KCNQ1	NM_000218.2(KCNQ1):c.409C>T (p.Leu137Phe)	CCTCATCGTCCTGGTCTGCYTCA	Congenital long QT syndrome, Cardiac arrhythmia, Long QT syndrome, LQT1 subtype
775479779	KCNQ1	NM_000218.2(KCNQ1):c.642C>A (p.Cys214Ter)	CCTCCATGGTGGTCCTCTGHGTG, CCATGGTGGTCCTCTGHGTGGGC	Cardiac arrhythmia
397508075	KCNQ1	NM_000218.2(KCNQ1):c.1075C>T (p.Gln359Ter)	CCTGAAGGTGCAGCAGAAGYAGA	Cardiac arrhythmia, Long QT syndrome, LQT1 subtype
397508097	KCNQ1	NM_000218.2(KCNQ1):c.1588C>T (p.Gln530Ter)	CCAAGAAGAAATTCYAGGTAAGC	Long QT syndrome, Cardiac arrhythmia
796052642	KCNQ2	NM_172107.2(KCNQ2):c.1009G>A (p.Ala337Thr)	CCCGGCARCAGGCCTGATCCAGG	not provided
397514581	KCNQ2	NM_172107.2(KCNQ2):c.638G>A (p.Arg213Gln)	TGGACCRGCGGGGAGGCACCTGG	Early infantile epileptic encephalopathy 7, not provided
397514582	KCNQ2	NM_172107.2(KCNQ2):c.869G>A (p.Gly290Asp)	CTGGAACGRCAGGCTCCTTGCGG	Early infantile epileptic encephalopathy 7
587777219	KCNQ2	NM_172107.2(KCNQ2):c.794C>T (p.Ala265Val)	CATCCRCGTAGGTGTCAAAGTGG	Early infantile epileptic encephalopathy 7, not provided
796052634	KCNQ2	NM_172107.2(KCNQ2):c.809G>A (p.Trp270Ter)	TGGTRGGGCCTGGTGAGTTGTGG	not provided
118192190	KCNQ2	NM_172107.2(KCNQ2):c.296+1G>A	CTACGTRTGAGTGGCCGGCGGGG, CCTACGTRTGAGTGGCCGGCGGG, GCCTACGTRTGAGTGGCCGGCGG	Benign familial neonatal seizures 1
118192200	KCNQ2	NM_172107.2(KCNQ2):c.620G>A (p.Arg207Gln)	TGCRGATGATCCGCATGGACCGG, GATTCTGCRGATGATCCGCATGG	Benign familial neonatal seizures 1, Seizures, benign familial neonatal, 1, and/or myokymia, not provided
118192216	KCNQ2	NM_172107.2(KCNQ2):c.998G>A (p.Arg333Gln)	AAGAGGCRGAACCCGGCAGCAGG	Benign familial neonatal seizures 1
794727740	KCNQ2	NM_172107.2(KCNQ2):c.793G>A (p.Ala265Thr)	TACRCGGATGCACTCTGGTGGGG, CTACRCGGATGCACTCTGGTGGG, CCTACRCGGATGCACTCTGGTGG, ACACCTACRCGGATGCACTCTGG	Benign familial neonatal seizures 1, Seizures, Early infantile epileptic encephalopathy 7, not provided
118192214	KCNQ2	NM_172107.2(KCNQ2):c.967C>T (p.Gln323Ter)	CCCTGAAGGTTCAGGAGYAGCAC, CCTGAAGGTTCAGGAGYAGCACA	Benign familial neonatal seizures 1
118192224	KCNQ2	NM_172107.2(KCNQ2):c.1288C>T (p.Pro430Ser)	CCCCTGTGTGGATGCTGCYCCGG, CCCTGTGTGGATGCTGCYCCGGA,	Benign familial neonatal seizures 1

			CCTGTGTGGATGCTGCYCCGGAC	
118192236	KCNQ2	NM_172107.2(KCNQ2):c.1741C>T (p.Arg581Ter)	CCACCTGGACATGCTGTCCYGAA, CCTGGACATGCTGTCCYGAATTA	Benign familial neonatal seizures 1, not provided
118192251	KCNQ3	NM_004519.3(KCNQ3):c.988C>T (p.Arg330Cys)	CCCAAAACGTGGGAAGGCYGTCT , CCAAAACGTGGGAAGGCYGTCTG	Benign familial neonatal seizures 2
28939710	KCNQ4	NM_004700.3(KCNQ4):c.961G>A (p.Gly321Ser)	AGGCTCCRGCTTTGCCCTGAAGG	DFNA 2 Nonsyndromic Hearing Loss
80358279	KCNQ4	NM_004700.3(KCNQ4):c.886G>A (p.Gly296Ser)	CTGRGCAGGGTCCTGGCTGCTGG	DFNA 2 Nonsyndromic Hearing Loss
397515402	KCNT1	NM_020822.2(KCNT1):c.1283G>A (p.Arg428Gln)	CAGCRGGTCATCTACCTCCAGGG, CCAGCRGGTCATCTACCTCCAGG	Early infantile epileptic encephalopathy 14, Epilepsy, nocturnal frontal lobe, 5
397515404	KCNT1	NM_020822.2(KCNT1):c.1421G>A (p.Arg474His)	CCTGCRCGCCTGGGCCGTGAAGG	Early infantile epileptic encephalopathy 14
587777264	KCNT1	NM_020822.2(KCNT1):c.862G>A (p.Gly288Ser)	GACCTGCRGCATCCAGCACCTGG	Early infantile epileptic encephalopathy 14
397515405	KCNT1	NM_020822.2(KCNT1):c.2782C>T (p.Arg928Cys)	CCCACCCTTCCAACATGYGCTTC, CCACCCTTCCAACATGYGCTTCA	Epilepsy, nocturnal frontal lobe, 5
387907302	KCNV2	NM_133497.3(KCNV2):c.226C>T (p.Gln76Ter)	CCTGGCAGAAGAGGACYAGCAGG	Retinal cone dystrophy 3B
786205860	KCTD17	NM_001282684.1(KCTD17):c.434G>A (p.Arg145His)	GTACCRCGTGCTGCAGTGCCAGG	DYSTONIA 26, MYOCLONIC
387907260	KCTD7	NM_153033.4(KCTD7):c.280C>T (p.Arg94Trp)	CCCCACGGACTCCGAGGGCYGGT , CCCACGGACTCCGAGGGCYGGTA , CCACGGACTCCGAGGGCYGGTAC	Epilepsy, progressive myoclonic 3
199422234	KDM5C	NM_004187.3(KDM5C):c.2191C>T (p.Leu731Phe)	CCCAGACGGCCTTGTCTGCYTTT, CCAGACGGCCTTGTCTGCYTTTC	Mental retardation, syndromic, Claes-Jensen type, X-linked
121917860	KERA	NM_007035.3(KERA):c.520C>T (p.Gln174Ter)	CCTGACCCTTCTTGACCTAYAGA, CCCTTCTTGACCTAYAGAACAAC	Cornea plana 2
730882122	KIF11	NM_004523.3(KIF11):c.790-1G>A	AAATTAAARGTTGATCTTGCAGG	Microcephaly with or without chorioretinopathy, lymphedema, or mental retardation
672601369	KIF1A	NM_001244008.1(KIF1A):c.757G>A (p.Glu253Lys)	AGCRAGCGGGCTGACTCCACGGG , GAGCRAGCGGGCTGACTCCACGG	Mental retardation, autosomal dominant 9
267607200	KIF21A	NM_001173464.1(KIF21A):c.2841G>A (p.Met947Ile)	AGATATRAATAGACTCCTCAAGG	Fibrosis of extraocular muscles, congenital, 3b
121912585	KIF21A	NM_001173464.1(KIF21A)	CCTGCTTTCTCAATAGCAAYGGG	Fibrosis of extraocular muscles, congenital, 1, Fibrosis of

		:c.2860C>T (p.Arg954Trp)		extraocular muscles, congenital, 3b
387907288	KIF5A	NM_004984.2(KIF5A):c.839G>A (p.Arg280His)	ATCRTGACAGCAAAATGACAAGG	Spastic paraplegia 10
387907289	KIF5A	NM_004984.2(KIF5A):c.704G>A (p.Gly235Glu)	GCAGRGAGTGAGAAGGTAGGGG, GGCAGRGAGTGAGAAGGTAGGGG, TGGCAGRGAGTGAGAAGGTAGGG, CTGGCAGRGAGTGAGAAGGTAGG	Spastic paraplegia 10
121434444	KIF5A	NM_004984.2(KIF5A):c.1082C>T (p.Ala361Val)	CCCAGAAGGAGACGATTGYGAAG, CCAGAAGGAGACGATTGYGAAGC	Spastic paraplegia 10
794727316	KIF7	NM_198525.2(KIF7):c.61C>T (p.Arg21Ter)	CCCAGTGCGGGTTGCCCTGYGAG, CCAGTGCGGGTTGCCCTGYGAGT	Acrocallosal syndrome, Schinzel type
104894701	KISS1R	NM_032551.4(KISS1R):c.991C>T (p.Arg331Ter)	CCTGGGCTCGCACTTCYGACAGG	
794726675	KIT	NM_000222.2(KIT):c.1879+1G>A	CTCAAGCRTAAGTTCCTGTATGG	Partial albinism
121913680	KIT	NM_000222.2(KIT):c.1747G>A (p.Glu583Lys)	AATGGRAGTTTCCCAGAAACAGG	Partial albinism
370756367	KLHL10	NM_152467.3(KLHL10):c.937G>A (p.Ala313Thr)	TGACRCTCGGGCAGACAGATGGG, ATGACRCTCGGGCAGACAGATGG	Spermatogenic failure 11
199469643	KLHL3	NM_017415.2(KLHL3):c.1292G>A (p.Arg431Gln)	CGGCRGAGCAGTGTGGGTGTGGG, GCGGCRGAGCAGTGTGGGTGTGG	Pseudohypoaldosteronism, type 2
199469628	KLHL3	NM_017415.2(KLHL3):c.1019C>T (p.Ala340Val)	CCTTCCAGAAGATGCAGAGYAGG, CCAGAAGATGCAGAGYAGGTGAG	Pseudohypoaldosteronism, type 2
730882260	KLHL41	NM_006063.2(KLHL41):c.1238C>T (p.Ser413Leu)	CCTTCAAACAGAGGCTTYGCTGG	Nemaline myopathy 9
104894704	KLK4	NM_004917.4(KLK4):c.458G>A (p.Trp153Ter)	GGCTRGGGTCTGCTGGCGAACGG	Amelogenesis imperfecta, hypomaturation type, IIA1
794727420	KMT2D	NM_003482.3(KMT2D):c.5	CCAAGGACCTGCAGYAGCTCTTC	Kabuki make-up syndrome

		677C>T (p.Gln1893Ter)		
794727549	KMT2D	NM_003482.3(KMT2D):c.7903C>T (p.Arg2635Ter)	CCCATGGAGCCTCACAGYGATCA, CCATGGAGCCTCACAGYGATCAG	Kabuki make-up syndrome
587783685	KMT2D	NM_003482.3(KMT2D):c.12592C>T (p.Arg4198Ter)	CCTACGGTGGGTCAGCTTYGAGC	Kabuki make-up syndrome
587783699	KMT2D	NM_003482.3(KMT2D):c.15943C>T (p.Gln5315Ter)	CCCGGGGTGGAGAGCTGTYAAAA, CCGGGGTGGAGAGCTGTYAAAAC	Kabuki make-up syndrome
587783711	KMT2D	NM_003482.3(KMT2D):c.3121C>T (p.Gln1041Ter)	CCAAAACTCCCCTCCTTCCYAGT	Kabuki make-up syndrome
398123704	KMT2D	NM_003482.3(KMT2D):c.11149C>T (p.Gln3717Ter)	CCTGATTCAAGGCTTTTAYAGGA	Kabuki make-up syndrome, not provided
398123708	KMT2D	NM_003482.3(KMT2D):c.11692C>T (p.Gln3898Ter)	CCCATGGGCTCTTTAYAGCAGCT, CCATGGGCTCTTTAYAGCAGCTT	Kabuki make-up syndrome, not provided
398123711	KMT2D	NM_003482.3(KMT2D):c.12406C>T (p.Gln4136Ter)	CCCCACCTGCTGGCTYAGCCCTC, CCCACCTGCTGGCTYAGCCCTCT	Kabuki make-up syndrome, not provided
398123721	KMT2D	NM_003482.3(KMT2D):c.14710C>T (p.Arg4904Ter)	CCAATCTGGATGTGYGACAGCTC	Kabuki make-up syndrome, not provided
398123757	KMT2D	NM_003482.3(KMT2D):c.7066C>T (p.Gln2356Ter)	CCAGGAGCCACCCCCTGCCYAGG	Kabuki make-up syndrome, not provided
59977263	KRT17	NM_000422.2(KRT17):c.304G>A (p.Val102Met)	CAAGRTGCGTGCCCTGGAGGAGG, GGACAAGRTGCGTGCCCTGGAGG	Pachyonychia congenita type 2, not provided
137852629	KRT2	NM_000423.2(KRT2):c.1459G>A (p.Glu487Lys)	AGGGCRAGGAGTGCAGGTGAGGG, GAGGGCRAGGAGTGCAGGTGAGG	Ichthyosis bullosa of Siemens, Ichthyosis exfoliativa, not provided
121912476	KRT5	NM_000424.3(KRT5):c.1252G>A (p.Glu418Lys)	TGCCRAGCAGCGTGGGGAGCTGG	Epidermolysis bullosa simplex, autosomal recessive, not provided
57499817	KRT5	NM_000424.3(KRT5):c.74C>T (p.Pro25Leu)	CCGCCTCTGCCATCACCCYGTCT, CCTCTGCCATCACCCYGTCTGTC	Epidermolysis bullosa simplex with mottled pigmentation, not provided
60554162	KRT6A	NM_005554.3(KRT6A):c.1414G>A (p.Glu472Lys)	GGAGGGTRAGGAGTGCAGGTGGG, TGGAGGGTRAGGAGTGCAGGTGG	PC-K6a, not provided
60627726	KRT6B	NM_005555.3(KRT6B):c.1414G>A (p.Glu472Lys)	GGAGGGCRAGGAGTGCAGGTGGG,	Pachyonychia congenita 4, not provided

			TGGAGGGCRAGGAGTGCAGGTG G	
587777292	KRT6C	NM_173086.4(KRT6C):c.1 414G>A (p.Glu472Lys)	CCTACCTGCACTCCTYGCCCTCC	Palmoplantar keratoderma, nonepidermolytic, focal or diffuse
57802288	KRT83	NM_002282.3(KRT83):c.12 19G>A (p.Glu407Lys)	ATATCRAGATCGCCACCTACAGG	Beaded hair, not provided
57019720	KRT9	NM_000226.3(KRT9):c.511 G>A (p.Val171Met)	TAAGRTGCAGGCTCTAGAGGAGG , GGATAAGRTGCAGGCTCTAGAGG	Epidermolytic palmoplantar keratoderma, not provided
137852519	L1CAM	NM_000425.4(L1CAM):c.1 792G>A (p.Asp598Asn)	ACTGRATGTGGTGGAGAGTAGGG , AACTGRATGTGGTGGAGAGTAGG	Spastic paraplegia 1
137852524	L1CAM	NM_000425.4(L1CAM):c.1 108G>A (p.Gly370Arg)	ATCAACRGGATCCCTGTGGAGGG , AATCAACRGGATCCCTGTGGAGG	Spastic paraplegia 1
137852525	L1CAM	NM_000425.4(L1CAM):c.2 254G>A (p.Val752Met)	CGCRTGCAGTGGCGCCCTCAGGG , CCGCRTGCAGTGGCGCCCTCAGG	
118204021	L2HGDH	NM_024884.2(L2HGDH):c. 164G>A (p.Gly55Asp)	TCGTTGRTGGCGGAATTGTGGGG, ATCGTTGRTGGCGGAATTGTGGG, CATCGTTGRTGGCGGAATTGTGG	L-2-hydroxyglutaric aciduria
121913574	LAMA2	NM_000426.3(LAMA2):c.1 580G>A (p.Cys527Tyr)	ACAGATRTCAGAGTTCCTACTGG	Congenital muscular dystrophy due to partial LAMA2 deficiency, not provided
398123367	LAMA2	NM_000426.3(LAMA2):c.1 12+1G>A	AAGAGRTACAGTCGAGGCATGG G, AAAGAGRTACAGTCGAGGCATG G	Merosin deficient congenital muscular dystrophy, not provided
9492297	LAMA2	NM_000426.3(LAMA2):c.2 750-1G>C	GCARCCTGTCGCTGTAATGCCGG	Merosin deficient congenital muscular dystrophy, not provided
398123391	LAMA2	NM_000426.3(LAMA2):c.9 212-1G>A	TCARATGACCTCAAGCAGTTTGG	Merosin deficient congenital muscular dystrophy, not provided
121913571	LAMA2	NM_000426.3(LAMA2):c.9 253C>T (p.Arg3085Ter)	CCAGTATTCCGTTCYGAGGTTGC	Merosin deficient congenital muscular dystrophy
121913572	LAMA2	NM_000426.3(LAMA2):c.7 732C>T (p.Arg2578Ter)	CCACCTAGGAGAAAAYGAAGGC A	Merosin deficient congenital muscular dystrophy, not provided
727502851	LAMA2	NM_000426.3(LAMA2):c.7 888C>T (p.Arg2630Ter)	CCGTTCATGTAGAGYGAACTAGA	Congenital muscular dystrophy
398123373	LAMA2	NM_000426.3(LAMA2):c.3 976C>T (p.Arg1326Ter)	CCATAGAACTGTGACCYGAGAAG	Merosin deficient congenital muscular dystrophy, not provided

398123378	LAMA2	NM_000426.3(LAMA2):c.5914C>T (p.Gln1972Ter)	CCAAAGGCTGTCTTYAGAAAAGC	Merosin deficient congenital muscular dystrophy, not provided
137852757	LAMA3	NM_198129.2(LAMA3):c.6808C>T (p.Arg2270Ter)	CCAATCTCACAACTCTCYGAGAT	Junctional epidermolysis bullosa gravis of Herlitz
137852758	LAMA3	NM_198129.2(LAMA3):c.8962C>T (p.Gln2988Ter)	CCAATCATGGAGCCCTCYAGTTT	Adult junctional epidermolysis bullosa
730880125	LAMB2	NM_002292.3(LAMB2):c.2890C>T (p.Arg964Ter)	CCGTGGGCAGGGCTGYGATGTGA	Pierson syndrome
267607208	LAMB2	NM_002292.3(LAMB2):c.4177C>T (p.Leu1393Phe)	CCAGCGGGCACTTGGCAAGYTCT	Nephrotic syndrome, type 5, with or without ocular abnormalities
121912484	LAMB3	NM_000228.2(LAMB3):c.1830G>A (p.Trp610Ter)	GTGRTCAGGGCCTGGGCTGGAGG, CCTGTGRTCAGGGCCTGGGCTGG	Junctional epidermolysis bullosa gravis of Herlitz
80356681	LAMB3	NM_000228.2(LAMB3):c.727C>T (p.Gln243Ter)	CCTACTATGCTGTGTCCYAGCTC	Junctional epidermolysis bullosa gravis of Herlitz
121912483	LAMB3	NM_000228.2(LAMB3):c.496C>T (p.Gln166Ter)	CCGCCAGGGTCGGCCTYAGAGCT	Junctional epidermolysis bullosa gravis of Herlitz
587776812	LAMB3	NM_000228.2(LAMB3):c.628+42G>A	CCAACTCTGTTTCCTTTYCCACC	Adult junctional epidermolysis bullosa
387906887	LAMC3	NM_006059.3(LAMC3):c.1156C>T (p.Gln386Ter)	CCAGCCCTGTGACTGCYAGTCGG	Cortical malformations, occipital
104894858	LAMP2	NM_002294.2(LAMP2):c.928G>A (p.Val310Ile)	GGCTCCRGTAAGCAAAGCACTGG	Cardiomyopathy, Danon disease
398124181	LARGE	NM_004737.4(LARGE):c.1102C>T (p.Gln368Ter)	CCACACCCGCTCCGAGYAGTGCT	not provided
398123036	LARS2	NM_015340.3(LARS2):c.1886C>T (p.Thr629Met)	CCTGTTCATGCAAAAAYGAAAGA	Perrault syndrome 4
121908050	LCAT	NM_000229.1(LCAT):c.440C>T (p.Thr147Ile)	CCCACAGGGTACCTGCACAYACT, CCACAGGGTACCTGCACAYACTG	Fish-eye disease
387906300	LCAT	NM_000229.1(LCAT):c.544C>T (p.Arg182Cys)	CCAGCAGGAGGAGTACTACYGCA	Norum disease
45514002	LDB3	NM_007078.2(LDB3):c.2017G>A (p.Asp673Asn)	TGCRATTTCCCCGTGGAGGCTGG, TGGCTGCRATTTCCCCGTGGAGG	Dilated cardiomyopathy 1C, Left ventricular noncompaction 3
45487699	LDB3	NM_007078.2(LDB3):c.566C>T (p.Ser189Leu)	CCAAAAGCCCTGCCGGGCTYGAG	Dilated cardiomyopathy 1C, not specified, not provided, Familial hypertrophic cardiomyopathy 24
121908335	LDB3	NM_001080116.1(LDB3):c.802C>T (p.Arg268Cys)	CCCACGTTTTGCCAAATTGYGCA, CCACGTTTTGCCAAATTGYGCAA	Myofibrillar myopathy, ZASP-related, not specified
121908337	LDB3	NM_007078.2(LDB3):c.617	CCTGTACTCGGCAGAGAYCCTGA	Dilated cardiomyopathy 1C

		C>T (p.Thr206Ile)		
375009082	LDLR	NM_000527.4(LDLR):c.2098G>A (p.Asp700Asn)	CTGCCCGRACGGCATGCTGCTGG	not provided
121908033	LDLR	NM_000527.4(LDLR):c.523G>A (p.Asp175Asn)	GCGAARATGGCTCGGATGAGTGG	Familial hypercholesterolemia
121908037	LDLR	NM_000527.4(LDLR):c.2531G>A (p.Gly844Asp)	GGACGRCTACAGCTACCCCTCGG	Familial hypercholesterolemia
768563000	LDLR	NM_000527.4(LDLR):c.718G>A (p.Glu240Lys)	GACRAATTCCAGTGCTCTGATGG	not provided
28942081	LDLR	NM_000527.4(LDLR):c.1637G>A (p.Gly546Asp)	GAAAGGGGRCCTGAATGGTGTGG	Familial hypercholesterolemia
139361635	LDLR	NM_000527.4(LDLR):c.1024G>A (p.Asp342Asn)	CCCCRACGGCTTCCAGCTGGTGG, GTGCCCCRACGGCTTCCAGCTGG	Hypercholesterolaemia, not provided
387906303	LDLR	NM_000527.4(LDLR):c.670G>A (p.Asp224Asn)	CTGCAAGRACAAATCTGACGAGG	Familial hypercholesterolemia
121908026	LDLR	NM_000527.4(LDLR):c.530C>T (p.Ser177Leu)	CCCGACTGCGAAGATGGCTYGGA, CCGACTGCGAAGATGGCTYGGAT	Familial hypercholesterolemia, not provided
121908044	LDLR	NM_000527.4(LDLR):c.621C>T (p.Gly207=)	CCACTGCCTAAGTGGYGAGTGCA	Familial hypercholesterolemia
752596535	LDLR	NM_000527.4(LDLR):c.501C>A (p.Cys167Ter)	CCCCCAGCTGTGGGCCTGHGACA, CCCCAGCTGTGGGCCTGHGACAA, CCCAGCTGTGGGCCTGHGACAAC, CCAGCTGTGGGCCTGHGACAACG	not provided
121908324	LDLRAP1	NM_015627.2(LDLRAP1):c.65G>A (p.Trp22Ter)	AGAGCTRGGGGGGCGGTGGCCGG	Hypercholesterolemia, autosomal recessive
121909126	LEFTY2	NM_003240.3(LEFTY2):c.1025G>A (p.Ser342Asn)	GGTCARCCTGCCCAACATGAGGG, TGGTCARCCTGCCCAACATGAGG	Left-right axis malformations
121909125	LEFTY2	NM_003240.3(LEFTY2):c.940C>T (p.Arg314Ter)	CCATTTCTGGGGCCGYGACAGTG	Left-right axis malformations
121912517	LHB	NM_000894.2(LHB):c.167G>A (p.Gly56Asp)	TGCCGRCTACTGCCCCACCATGG	Isolated lutropin deficiency
137854503	LHX3	NM_178138.4(LHX3):c.629C>T (p.Ala210Val)	CCAGAACCGCCGGGYCAAGGAGA	Pituitary hormone deficiency, combined 3
121912642	LHX4	NM_033343.3(LHX4):c.25	CCTGCTGCCCTGACAGGYGCTTC	Pituitary hormone deficiency, combined 4