[0007] Therefore, the development of programmable editors that are flexibly capable of introducing any desired single nucleotide change and/or which could install base pair insertions or deletions (e.g., at least 1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, or more base pair insertions or deletions) and/or which could alter or modify the nucleotide sequence at a target site with high specificity and efficiency would substantially expand the scope and therapeutic potential of genome editing technologies based on CRISPR.

SUMMARY OF THE INVENTION

[0008] The present disclosure describes an entirely new platform for genome editing called "prime editing." Prime editing is a versatile and precise genome editing method that directly writes new genetic information into a specified DNA site using a nucleic acid programmable DNA binding protein ("napDNAbp") working in association with a polymerase (i.e., in the form of a fusion protein or otherwise provided *in trans* with the napDNAbp), wherein the prime editing system is programmed with a prime editing (PE) guide RNA ("PEgRNA") that both specifies the target site and templates the synthesis of the desired edit in the form of a replacement DNA strand by way of an extension (either DNA or RNA) engineered onto a guide RNA (e.g., at the 5′ or 3′ end, or at an internal portion of a guide RNA). The replacement strand containing the desired edit (e.g., a single nucleobase substitution) shares the same sequence as the endogenous strand of the target site to be edited (with the exception that it includes the desired edit). Through DNA repair and/or replication machinery, the endogenous strand of the target site is replaced by the newly synthesized replacement strand containing the desired edit. In some cases, prime editing may be thought of as a "search-and-replace" genome editing technology since the prime editors, as described herein, not only search and locate the desired target site to be edited, but at the same time, encode a replacement strand containing a desired edit which is installed in place of the corresponding target site endogenous DNA strand.

[0009] The prime editors of the present disclosure relate, in part, to the discovery that the mechanism of target-primed reverse transcription (TPRT) or "prime editing" can be leveraged or adapted for conducting precision CRISPR/Cas-based genome editing with high efficiency and genetic flexibility (e.g., as depicted in various embodiments of FIGs. 1A-1F). TPRT is naturally used by mobile DNA elements, such as mammalian non-LTR retrotransposons and bacterial Group II introns[28,29]. The inventors have herein used Cas protein-reverse transcriptase fusions or related systems to target a specific DNA sequence

with a guide RNA, generate a single strand nick at the target site, and use the nicked DNA as a primer for reverse transcription of an engineered reverse transcriptase template that is integrated with the guide RNA. However, while the concept begins with prime editors that use reverse transcriptases as the DNA polymerase component, the prime editors described herein are not limited to reverse transcriptases but may include the use of virtually and DNA polymerase. Indeed, while the application throughout may refer to prime editors with "reverse transcriptases," it is set forth here that reverse transcriptases are only one type of DNA polymerase that may work with prime editing. Thus, wherever the specification mentions "reverse transcriptases," the person having ordinary skill in the art should appreciate that any suitable DNA polymerase may be used in place of the reverse transcriptase. Thus, in one example, the prime editors may comprise Cas9 (or an equivalent napDNAbp) which is programmed to target a DNA sequence by associating it with a specialized guide RNA (i.e., PEgRNA) containing a spacer sequence that anneals to a complementary protospacer in the target DNA. The specialized guide RNA also contains new genetic information in the form of an extension that encodes a replacement strand of DNA containing a desired genetic alteration which is used to replace a corresponding endogenous DNA strand at the target site. To transfer information from the PEgRNA to the target DNA, the mechanism of prime editing involves nicking the target site in one strand of the DNA to expose a 3′-hydroxyl group. The exposed 3′-hydroxyl group can then be used to prime the DNA polymerization of the edit-encoding extension on PEgRNA directly into the target site. In various examples, the extension—which provides the template for polymerization of the replacement strand containing the edit—can be formed from RNA or DNA. In the case of an RNA extension, the polymerase of the prime editor can be an RNA-dependent DNA polymerase (such as, a reverse transcriptase). In the case of a DNA extension, the polymerase of the prime editor may be a DNA-dependent DNA polymerase.

[0010] The newly synthesized strand (i.e., the replacement DNA strand containing the desired edit) that is formed by the herein disclosed prime editors would be homologous to the genomic target sequence (i.e., have the same sequence as) except for the inclusion of a desired nucleotide change (e.g., a single nucleotide change, a deletion, or an insertion, or a combination thereof). The newly synthesized (or replacement) strand of DNA may also be referred to as a single strand DNA flap, which would compete for hybridization with the complementary homologous endogenous DNA strand, thereby displacing the corresponding endogenous strand. In certain examples, the system can be combined with the use of an error-prone reverse transcriptase enzyme (e.g., provided as a fusion protein with the Cas9

domain, or provided *in trans* to the Cas9 domain). The error-prone reverse transcriptase enzyme can introduce alterations during synthesis of the single strand DNA flap. Thus, in certain examples, error-prone reverse transcriptase can be utilized to introduce nucleotide changes to the target DNA. Depending on the error-prone reverse transcriptase that is used with the system, the changes can be random or non-random.

[0011] Resolution of the hybridized intermediate (comprising the single strand DNA flap synthesized by the reverse transcriptase hybridized to the endogenous DNA strand) can include removal of the resulting displaced flap of endogenous DNA (e.g., with a 5′ end DNA flap endonuclease, FEN1), ligation of the synthesized single strand DNA flap to the target DNA, and assimilation of the desired nucleotide change as a result of cellular DNA repair and/or replication processes. Because templated DNA synthesis offers single nucleotide precision for the modification of any nucleotide, including insertions and deletions, the scope of this approach is very broad and could foreseeably be used for myriad applications in basic science and therapeutics.

[0012] In one example, the specification provides a fusion protein comprising a nucleic acid programmable DNA binding protein (napDNAbp) and a reverse transcriptase. In various examples, the fusion protein is capable of carrying out genome editing by target-primed reverse transcription in the presence of an extended guide RNA.

[0013] In certain examples, the napDNAbp has a nickase activity. The napDNAbp may also be a Cas9 protein or functional equivalent thereof, such as a nuclease active Cas9, a nuclease inactive Cas9 (dCas9), or a Cas9 nickase (nCas9).

[0014] In certain examples, the napDNAbp is selected from the group consisting of: Cas9, Cas12e, Cas12d, Cas12a, Cas12b1, Cas13a, Cas12c, and Argonaute and optionally has a nickase activity.

[0015] In other examples, the fusion protein when complexed with an extended guide RNA is capable of binding to a target DNA sequence.

[0016] In still other examples, the target DNA sequence comprises a target strand and a complementary non-target strand.

[0017] In other examples, the binding of the fusion protein complexed to the extended guide RNA forms an R-loop. The R-loop can comprise (i) an RNA-DNA hybrid comprising the extended guide RNA and the target strand, and (ii) the complementary non-target strand.

[0018] In still other examples, the complementary non-target strand is nicked to form a reverse transcriptase priming sequence having a free 3′ end.

[0019] In various examples, the extended guide RNA comprises (a) a guide RNA and (b) an RNA extension at the 5′ or the 3′ end of the guide RNA, or at an intramolecular location in the guide RNA. The RNA extension can comprise (i) a reverse transcription template sequence comprising a desired nucleotide change, (ii) a reverse transcription primer binding site, and (iii) optionally, a linker sequence. In various examples, the reverse transcription template sequence may encode a single-strand DNA flap that is complementary to an endogenous DNA sequence adjacent to the nick site, wherein the single-strand DNA flap comprises the desired nucleotide change.

[0020] In various examples, the RNA extension is at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 21 nucleotides, at least 22 nucleotides, at least 23 nucleotides, at least 24 nucleotides, or at least 25 nucleotides in length.

[0021] In still other examples, the single-strand DNA flap may hybridize to the endogenous DNA sequence adjacent to the nick site, thereby installing the desired nucleotide change. In still other examples, the single-stranded DNA flap displaces the endogenous DNA sequence adjacent to the nick site and which has a free 5′ end. In certain examples, the displaced endogenous DNA having the 5′ end is excised by the cell.

[0022] In various examples, the cellular repair of the single-strand DNA flap results in installation of the desired nucleotide change, thereby forming a desired product.

[0023] In various other examples, the desired nucleotide change is installed in an editing window that is between about -4 to +10 of the PAM sequence.

[0024] In still other examples, the desired nucleotide change is installed in an editing window that is between about -5 to +5 of the nick site, or between about -10 to +10 of the nick site, or between about -20 to +20 of the nick site, or between about -30 to +30 of the nick site, or between about -40 to + 40 of the nick site, or between about -50 to +50 of the nick site, or between about -60 to +60 of the nick site, or between about -70 to +70 of the nick site, or between about -80 to +80 of the nick site, or between about -90 to +90 of the nick site, or between about -100 to +100 of the nick site, or between about -200 to +200 of the nick site.

[0025] In various examples, the napDNAbp comprises an amino acid sequence of SEQ ID NO: 18. In various other examples, the napDNAbp comprises an amino acid

sequence that is at least 80%, 85%, 90%, 95%, 98%, or 99% identical to the amino acid sequence of any one of SEQ ID NOs: 26-39, 42- 61, 75-76, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, and 482-487 (Cas9); (SpCas9); SEQ ID NO: 77-86 (CP-Cas9); SEQ ID NO: 18-25 and 87-88 (SpCas9); and SEQ ID NOs: 62-72(Cas12)

[0026] In other examples, the reverse transcriptase of the disclosed fusion proteins and/or compositions may comprise any one of the amino acid sequences of SEQ ID NOs: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700-716, 739-742, and 766. In still other examples, the reverse transcriptase may comprise an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, or 99% identical to the amino acid sequence of any one of SEQ ID NOs: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700-716, 739-742, and 766. These sequences may be naturally occurring reverse transcriptase sequences, e.g., from a retrovirus or a retrotransposon, of the sequences may be recombinant.

[0027] In various other examples, the fusion proteins herein disclosed may comprise various structural configurations. For example, the fusion proteins may comprise the structure NH_2-[napDNAbp]-[reverse transcriptase]-COOH; or NH_2-[reverse transcriptase]-[napDNAbp]-COOH, wherein each instance of "]-["indicates the presence of an optional linker sequence.

[0028] In various examples, the linker sequence comprises an amino acid sequence of SEQ ID NOs: 127, 165-176, 446, 453, and 767-769, or an amino acid sequence that this at least 80%, 85%, or 90%, or 95%, or 99% identical to any one of the linker amino acid sequence of SEQ ID NOs: 127, 165-176, 446, 453, and 767-769.

[0029] In various examples, the desired nucleotide change that is incorporated into the target DNA can be a single nucleotide change (e.g., a transition or transversion), an insertion of one or more nucleotides, or a deletion of one or more nucleotides.

[0030] In certain cases, the insertion is at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 200, at least 300, at least 400, or at least 500 nucleotides in length.

[0031] In certain other cases, the deletion is at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 200, at least 300, at least 400, or at least 500 nucleotides in length.

[0032] In another example, the present disclosure provides an extended guide RNA comprising a guide RNA and at least one RNA extension. The RNA extension can be positioned at the 3′ end of the guide RNA. In other examples, the RNA extension can be positioned at the 5′ of the guide RNA. In still other examples, the RNA extension can be positioned at an intramolecular position within the guide RNA, however, preferable, the intramolecular positioning of the extended portion does not disrupt the functioning of the protospacer.

[0033] In various examples, the extended guide RNA is capable of binding to a napDNAbp and directing the napDNAbp to a target DNA sequence. The target DNA sequence can comprise a target strand and a complementary non-target strand, wherein the guide RNA hybridizes to the target strand to form an RNA-DNA hybrid and an R-loop.

[0034] In various examples of the extended guide RNA, the at least one RNA extension can comprise a reverse transcription template sequence. In various other example, the RNA extension may further comprises a reverse transcription primer binding site. In still further examples, the RNA extension may comprise a linker or spacer sequence that joins the RNA extension to the guide RNA.

[0035] In various examples, the RNA extension can be at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 21 nucleotides, at least 22 nucleotides, at least 23 nucleotides, at least 24 nucleotides, at least 25 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90 nucleotides, at least 100 nucleotides, at least 150 nucleotides, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length.

[0036] In other examples, the reverse transcription template sequence is at least 3 nucleotides, at least 4 nucleotides, at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90

nucleotides, at least 100 nucleotides, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length.

[0037] In still other examples, wherein the reverse transcription primer binding site sequence is at least 3 nucleotides, at least 4 nucleotides, at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90 nucleotides, at least 100 nucleotides, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length.

[0038] In other examples, the optional linker or spacer sequence is at least 3 nucleotides, at least 4 nucleotides, at least 5 nucleotides, at least 6 nucleotides, at least 7 nucleotides, at least 8 nucleotides, at least 9 nucleotides, at least 10 nucleotides, at least 11 nucleotides, at least 12 nucleotides, at least 13 nucleotides, at least 14 nucleotides, at least 15 nucleotides, at least 16 nucleotides, at least 17 nucleotides, at least 18 nucleotides, at least 19 nucleotides, at least 20 nucleotides, at least 30 nucleotides, at least 40 nucleotides, at least 50 nucleotides, at least 60 nucleotides, at least 70 nucleotides, at least 80 nucleotides, at least 90 nucleotides, at least 100 nucleotides, at least 200 nucleotides, at least 300 nucleotides, at least 400 nucleotides, or at least 500 nucleotides in length.

[0039] In various examples of the extended guide RNAs, the reverse transcription template sequence may encode a single-strand DNA flap that is complementary to an endogenous DNA sequence adjacent to a nick site, wherein the single-strand DNA flap comprises a desired nucleotide change. The single-stranded DNA flap may displace an endogenous single-strand DNA at the nick site. The displaced endogenous single-strand DNA at the nick site can have a 5′ end and form an endogenous flap, which can be excised by the cell. In various examples, excision of the 5′ end endogenous flap can help drive product formation since removing the 5′ end endogenous flap encourages hybridization of the single-strand 3′ DNA flap to the corresponding complementary DNA strand, and the incorporation or assimilation of the desired nucleotide change carried by the single-strand 3′ DNA flap into the target DNA.

[0040] In various examples of the extended guide RNAs, the cellular repair of the single-strand DNA flap results in installation of the desired nucleotide change, thereby forming a desired product.

[0041] In certain examples, the PEgRNA comprises the nucleotide sequence of SEQ ID NOs: 101-104, 131, 181-183, 222-234, 237-244, 277, 324-330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 394, 429-442, 499-505, 641-649, 678-692, 735-736, 738, 757-761, 776-777, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3540, 3549-3556, 3628-3698, 3755-3810, 3874, 3890-3901, 3905-3911, 3913-3929, and 3972-3989, or a nucleotide sequence having at least 85%, or at least 90%, or at least 95%, or at least 98%, or at least 99% sequence identity with any one of SEQ ID NOs: 101-104, 181-183, 223-234, 237-244, 277, 324-330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 394, 429-442, 499-505, 641-649, 678-692, 735-736, 757-761, 776-777, 2997-3103, 3113-3121, 3305-3455, 3479-3493, 3522-3540, 3549-3556, 3628-3698, 3755-3810, 3874, 3890-3901, 3905-3911, 3913-3929, and 3972-3989.

[0042] In yet another example of the disclosure, the specification provides for complexes comprising a fusion protein described herein and any extended guide RNA described above.

[0043] In still other examples of the disclosure, the specification provides a complex comprising a napDNAbp and an extended guide RNA. The napDNAbp can be a Cas9 nickase, or can be an amino acid sequence of SEQ ID NOs: 42-57 (Cas9 nickase) and 65 (AsCas12a nickase), or an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, or 99% identical to the amino acid sequence of any one of SEQ ID NOs: 42-57 (Cas9 nickase) and 65 (AsCas12a nickase).

[0044] In various examples involving a complex, the extended guide RNA is capable of directing the napDNAbp to a target DNA sequence. In various examples, a reverse transcriptase may be provided *in trans*, i.e., provided from a different source than the complex itself. For example, a reverse transcriptase could be provided to the same cell having the complex by introducing a separate vector separately encoding the reverse transcriptase.

[0045] In yet another example, the specification provides polynucleotides. In certain examples, the polynucleotides may encode any of the fusion proteins disclosed herein. In certain other examples, the polynucleotides may encode any of the napDNAbps disclosed herein. In still further examples, the polynucleotides may encode any of the reverse transcriptases disclosed herein. In yet other examples, the polynucleotides may encode any of the extended guide RNAs disclosed herein, any of the reverse transcription template sequences, or any of the reverse transcription primer sites, or any of the optional linker sequences.

[0046] In still other examples, the specification provides vectors comprising the polynucleotides described herein. Thus, in certain examples, the vectors comprise polynucleotides for encoding the fusion proteins comprising a napDNAbp and a reverse transcriptase. In other examples, the vectors comprise polynucleotides that separately encode a napDNAbp and reverse transcriptase. In still other examples, the vectors may comprise polynucleotides that encode the extended guide RNAs. In various examples, the vectors may comprise one or more polynucleotides that encode napDNAbps, reverse transcriptase, and extended guide RNAs on the same or separate vectors.

[0047] In still other examples, the specification provides cells comprising a fusion protein as described herein and an extended guide RNA. The cells may be transformed with the vectors comprising the fusion proteins, napDNAbps, reverse transcriptase, and extended guide RNAs. These genetic elements may be comprised on the same vectors or on different vectors.

[0048] In yet another example, the specification provides pharmaceutical compositions. In certain examples, the pharmaceutical compositions comprise one or more of a napDNAbp, a fusion protein, a reverse transcriptase, and an extended guide RNA. In certain examples, the fusion protein described herein and a pharmaceutically acceptable excipient. In other examples, the pharmaceutical compositions comprise any extend guide RNA described herein and a pharmaceutically acceptable excipient. In still other examples, the pharmaceutical compositions comprise any extend guide RNA described herein in combination with any fusion protein described herein and a pharmaceutically acceptable excipient. In yet other examples, the pharmaceutical compositions comprise any polynucleotide sequence encoding one or more of a napDNAbp, a fusion protein, a reverse transcriptase, and an extended guide RNA. In still other examples, the various components disclosed herein may be separated into one or more pharmaceutical compositions. For example, a first pharmaceutical composition may comprise a fusion protein or a napDNAbp, a second pharmaceutical compositions may comprise a reverse transcriptase, and a third pharmaceutical composition may comprise an extended guide RNA.

[0049] In still a further example, the present disclosure provides kits. In one example, the kit comprises one or more polynucleotides encoding one or more components, including a fusion protein, a napDNAbp, a reverse transcriptase, and an extended guide RNA. The kits may also comprise vectors, cells, and isolated preparations of polypeptides, including any fusion protein, napDNAbp, or reverse transcriptase disclosed herein.

[0050] In yet another example, the present disclosure provides for methods of using the disclosed compositions of matter.

[0051] In one example, the methods relate to a method for installing a desired nucleotide change in a double-stranded DNA sequence. The method first comprises contacting the double-stranded DNA sequence with a complex comprising a fusion protein and an extended guide RNA, wherein the fusion protein comprises a napDNAbp and a reverse transcriptase and wherein the extended guide RNA comprises a reverse transcription template sequence comprising the desired nucleotide change. Next, the method involves nicking the double-stranded DNA sequence on the non-target strand, thereby generating a free single-strand DNA having a 3′ end. The method then involves hybridizing the 3` end of the free single-strand DNA to the reverse transcription template sequence, thereby priming the reverse transcriptase domain. The method then involves polymerizing a strand of DNA from the 3′ end, thereby generating a single-strand DNA flap comprising the desired nucleotide change. Then, the method involves replacing an endogenous DNA strand adjacent the cut site with the single-strand DNA flap, thereby installing the desired nucleotide change in the double-stranded DNA sequence.

[0052] In other examples, the disclosure provides for a method for introducing one or more changes in the nucleotide sequence of a DNA molecule at a target locus, comprising contacting the DNA molecule with a nucleic acid programmable DNA binding protein (napDNAbp) and a guide RNA which targets the napDNAbp to the target locus, wherein the guide RNA comprises a reverse transcriptase (RT) template sequence comprising at least one desired nucleotide change. Next, the method involves forming an exposed 3′ end in a DNA strand at the target locus and then hybridizing the exposed 3′ end to the RT template sequence to prime reverse transcription. Next, a single strand DNA flap comprising the at least one desired nucleotide change based on the RT template sequence is synthesized or polymerized by reverse transcriptase. Lastly, the at least one desired nucleotide change is incorporated into the corresponding endogenous DNA, thereby introducing one or more changes in the nucleotide sequence of the DNA molecule at the target locus.

[0053] In still other examples, the disclosure provides a method for introducing one or more changes in the nucleotide sequence of a DNA molecule at a target locus by target-primed reverse transcription, the method comprising: (a) contacting the DNA molecule at the target locus with a (i) fusion protein comprising a nucleic acid programmable DNA binding protein (napDNAbp) and a reverse transcriptase and (ii) a guide RNA comprising an RT template comprising a desired nucleotide change; (b) conducting target-primed reverse transcription of

the RT template to generate a single strand DNA comprising the desired nucleotide change; and (c) incorporating the desired nucleotide change into the DNA molecule at the target locus through a DNA repair and/or replication process.

[0054] In certain examples, the step of replacing the endogenous DNA strand comprises: (i) hybridizing the single-strand DNA flap to the endogenous DNA strand adjacent the cut site to create a sequence mismatch; (ii) excising the endogenous DNA strand; and (iii) repairing the mismatch to form the desired product comprising the desired nucleotide change in both strands of DNA.

[0055] In various examples, the desired nucleotide change can be a single nucleotide substitution (e.g., and transition or a transversion change), a deletion, or an insertion. For example, the desired nucleotide change can be (1) a G to T substitution, (2) a G to A substitution, (3) a G to C substitution, (4) a T to G substitution, (5) a T to A substitution, (6) a T to C substitution, (7) a C to G substitution, (8) a C to T substitution, (9) a C to A substitution, (10) an A to T substitution, (11) an A to G substitution, or (12) an A to C substitution.

[0056] In other examples, the desired nucleoid change can convert (1) a G:C basepair to a T:A basepair, (2) a G:C basepair to an A:T basepair, (3) a G:C basepair to C:G basepair, (4) a T:A basepair to a G:C basepair, (5) a T:A basepair to an A:T basepair, (6) a T:A basepair to a C:G basepair, (7) a C:G basepair to a G:C basepair, (8) a C:G basepair to a T:A basepair, (9) a C:G basepair to an A:T basepair, (10) an A:T basepair to a T:A basepair, (11) an A:T basepair to a G:C basepair, or (12) an A:T basepair to a C:G basepair.

[0057] In still other examples, the method introduces a desired nucleotide change that is an insertion. In certain cases, the insertion is at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 200, at least 300, at least 400, or at least 500 nucleotides in length.

[0058] In other examples, the method introduces a desired nucleotide change that is a deletion. In certain other cases, the deletion is at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 200, at least 300, at least 400, or at least 500 nucleotides in length.

[0059] In various examples, the desired nucleotide change corrects a disease-associated gene. The disease-associated gene can be associated with a monogenetic disorder selected from the group consisting of: Adenosine Deaminase (ADA) Deficiency; Alpha-1 Antitrypsin Deficiency; Cystic Fibrosis; Duchenne Muscular Dystrophy; Galactosemia; Hemochromatosis; Huntington's Disease; Maple Syrup Urine Disease; Marfan Syndrome; Neurofibromatosis Type 1; Pachyonychia Congenita; Phenylkeotnuria; Severe Combined Immunodeficiency; Sickle Cell Disease; Smith-Lemli-Opitz Syndrome; and Tay-Sachs Disease. In other examples, the disease-associated gene can be associated with a polygenic disorder selected from the group consisting of: heart disease; high blood pressure; Alzheimer's disease; arthritis; diabetes; cancer; and obesity.

[0060] The methods disclosed herein may involve fusion proteins having a napDNAbp that is a nuclease dead Cas9 (dCas9), a Cas9 nickase (nCas9), or a nuclease active Cas9. In other examples, a napDNAbp and reverse transcriptase are not encoded as a single fusion protein, but rather can be provided in separate constructs. Thus, in some examples, the reverse transcriptase can be provided *in trans* relative to the napDNAbp (rather than by way of a fusion protein).

[0061] In various examples involving methods, the napDNAbp may comprise an amino acid sequence of SEQ ID NOs: 26- 61, 75-76, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, and 482-487 (Cas9); (SpCas9); SEQ ID NO: 77-86 (CP-Cas9); SEQ ID NO: 18 -25 and 87-88 (SpCas9); and SEQ ID NOs: 62-72(Cas12). The napDNAbp may also comprise an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, or 99% identical to the amino acid sequence of any one of SEQ ID NOs: 26- 61, 75-76, 126, 130, 137, 141, 147, 153, 157, 445, 460, 467, and 482-487 (Cas9); (SpCas9); SEQ ID NO: 77-86 (CP-Cas9); SEQ ID NO: 18 -25 and 87-88 (SpCas9); and SEQ ID NOs: 62-72(Cas12).

[0062] In various examples involving methods, the reverse transcriptase may comprise any one of the amino acid sequences of SEQ ID NOs: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700-716, 739-742, and 766. The reverse transcriptase may also comprise an amino acid sequence that is at least 80%, 85%, 90%, 95%, 98%, or 99% identical to the amino acid sequence of any one of SEQ ID NOs: 89-100, 105-122, 128-129, 132, 139, 143, 149, 154, 159, 235, 454, 471, 516, 662, 700-716, 739-742, and 766.

[0063] The methods may involve the use of a PEgRNA comprising a nucleotide sequence of SEQ ID NOs: 101-104, 131, 181-183, 222-234, 237-244, 277, 324-330, 332, 334, 336, 338, 340, 342, 344, 346, 348, 350, 352, 354, 356, 358, 360, 362, 364, 366, 368, 394, 429-442, 499-505, 641-649, 678-692, 735-736, 738, 757-761, 776-777, 2997-3103, 3113-3121, 3305-

3455, 3479-3493, 3522-3540, 3549-3556, 3628-3698, 3755-3810, 3874, 3890-3901, 3905-3911, 3913-3929, and 3972-3989, or a nucleotide sequence having at least 80%, or at least 85%, or at least 90%, or at least 95%, or at least 99% sequence identity thereto. The methods may comprise the use of extended guide RNAs that comprise an RNA extension at the 3′ end, wherein the RNA extension comprises the reverse transcription template sequence.

[0064] The methods may comprise the use of extended guide RNAs that comprise an RNA extension at the 5′ end, wherein the RNA extension comprises the reverse transcription template sequence.

[0065] The methods may comprise the use of extended guide RNAs that comprise an RNA extension at an intramolecular location in the guide RNA, wherein the RNA extension comprises the reverse transcription template sequence.

[0066] The methods may comprise the use of extended guide RNAs having one or more RNA extensions that are at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 30, at least 40, at least 50, at least 60, at least 70, at least 80, at least 90, at least 100, at least 200, at least 300, at least 400, or at least 500 nucleotides in length.

[0066a] In one aspect, the present disclosure provides an *in vitro* or *ex vivo* method for site-specific modification of a double-stranded target DNA sequence, the method comprising:

contacting the double-stranded target DNA sequence, which comprises a first strand and a second strand, with a prime editing system, wherein the prime editing system comprises:

(i) a fusion protein comprising a nucleic acid programmable DNA binding protein (napDNAbp) and a reverse transcriptase, wherein the napDNAbp is a Cas9 nickase, and

(ii) a prime editing guide RNA (PEgRNA) comprising:

(a) a spacer sequence that comprises a region of complementarity to the first strand of the double-stranded target DNA sequence;

(b) an extension arm that comprises a DNA synthesis template and a primer binding site in a 5′ to 3′ orientation, wherein the primer binding site comprises a region of complementarity to a region upstream of a nick site in the second strand of the double-stranded target DNA sequence, and wherein the DNA synthesis template encodes one or more nucleotide changes compared to a region downstream of the nick site in the second strand of the double-stranded target DNA sequence, and

(c) a gRNA core that interacts with the napDNAbp; wherein the contacting occurs in a cell and results in:

nicking the second strand of the double-stranded target DNA sequence to form a free 3′ end at the nick site;

annealing the primer binding site with the region of the second strand of the double-stranded target DNA sequence upstream of the nick site;

synthesizing a single strand of DNA encoded by the DNA synthesis template from the free 3′ end of the second strand of the double-stranded target DNA sequence by the reverse transcriptase; and

replacing the region downstream of the nick site in the second strand of the double-stranded target DNA sequence with the single strand of DNA, thereby modifying the sequence of the double-stranded target DNA sequence.

[0067] It should be appreciated that the foregoing concepts, and additional concepts discussed below, may be arranged in any suitable combination, as the present disclosure is not limited in this respect. Further, other advantages and novel features of the present disclosure will become apparent from the following detailed description of various non-limiting embodiments when considered in conjunction with the accompanying figures.

BRIEF DESCRIPTION OF THE DRAWINGS

[0068] The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure, which can be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

[0069] **FIG. 1A** provides a schematic of an exemplary process for introducing a single nucleotide change, and/or insertion, and/or deletion into a DNA molecule (e.g., a genome) using a fusion protein comprising a reverse transcriptase fused to a Cas9 protein in complex with an extended guide RNA molecule. In this embodiment, the guide RNA is extended at the 3′ end to include a reverse transcriptase template sequence. The schematic shows how a

CLAIMS

1. An *in vitro* or *ex vivo* method for site-specific modification of a double-stranded target DNA sequence, the method comprising:

contacting the double-stranded target DNA sequence, which comprises a first strand and a second strand, with a prime editing system, wherein the prime editing system comprises:

(i) a fusion protein comprising a nucleic acid programmable DNA binding protein (napDNAbp) and a reverse transcriptase, wherein the napDNAbp is a Cas9 nickase, and

(ii) a prime editing guide RNA (PEgRNA) comprising:

(a) a spacer sequence that comprises a region of complementarity to the first strand of the double-stranded target DNA sequence;

(b) an extension arm that comprises a DNA synthesis template and a primer binding site in a 5′ to 3′ orientation, wherein the primer binding site comprises a region of complementarity to a region upstream of a nick site in the second strand of the double-stranded target DNA sequence, and wherein the DNA synthesis template encodes one or more nucleotide changes compared to a region downstream of the nick site in the second strand of the double-stranded target DNA sequence, and

(c) a gRNA core that interacts with the napDNAbp; wherein the contacting occurs in a cell and results in:

nicking the second strand of the double-stranded target DNA sequence to form a free 3′ end at the nick site;

annealing the primer binding site with the region of the second strand of the double-stranded target DNA sequence upstream of the nick site;

synthesizing a single strand of DNA encoded by the DNA synthesis template from the free 3′ end of the second strand of the double-stranded target DNA sequence by the reverse transcriptase; and

replacing the region downstream of the nick site in the second strand of the double-stranded target DNA sequence with the single strand of DNA, thereby modifying the sequence of the double-stranded target DNA sequence.

2. The method of claim 1, wherein the one or more nucleotide changes are 1-150 bases downstream of the nick site in the second strand.

3. The method of claim 1 or 2, wherein the DNA synthesis template comprises a region of complementarity to the second strand of the double-stranded target DNA sequence.

4. The method of any one of claims 1-3, wherein the one or more nucleotide changes comprise a nucleotide substitution as compared to the region of the second strand downstream of the nick site in the double-stranded target DNA sequence.

5. The method of any one of claims 1-4, wherein the one or more nucleotide changes comprise a nucleotide deletion as compared to the region of the second strand downstream of the nick site in the double-stranded target DNA sequence.

6. The method of any one of claims 1-5, wherein the one or more nucleotide changes comprise a nucleotide insertion as compared to the region of the second strand downstream of the nick site in the double-stranded target DNA sequence.

7. The method of any one of claims 1-6, wherein the one or more nucleotide changes comprise two or more nucleotide insertions, nucleotide substitutions, nucleotide deletions, or a combination of two or more of a nucleotide insertion, a nucleotide substitution, or a nucleotide deletion, as compared to the region of the second strand downstream of the nick site in the double-stranded target DNA sequence.

8. The method of any one of claims 1-7, wherein the Cas9 nickase comprises a substitution in the HNH domain.

9. The method of claim 8, wherein the substitution in the HNH domain is H840A relative to SEQ ID NO: 18.

10. The method of any one of claims 1-9, wherein the reverse transcriptase is a retrovirus reverse transcriptase.

11. The method of any one of claims 1-10, wherein the reverse transcriptase is a Moloney Murine Leukemia reverse transcriptase (M-MLV RT).

12. The method of any one of claims 1-11 further comprising contacting the double-stranded target DNA sequence with a guide RNA, wherein the guide RNA directs the napDNAbp to generate a second nick in the first strand of the double-stranded target DNA sequence.

13. The method of claim 12, wherein the first nick and the second nick are at least 14 nucleotides apart.

14. The method of any one of claims 1-6 or 8-13, wherein the one or more nucleotide changes comprise a single nucleotide substitution, a single nucleotide insertion, a single nucleotide deletion, or combinations thereof, as compared to the region of the second strand downstream of the nick site in the double-stranded target DNA sequence.

15. The method of any one of claims 1-14, wherein the one or more nucleotide changes are 4-6 bases downstream of the position of the nick site in the second strand in the double-stranded target DNA sequence.

16. The method of any one of claims 1-14, wherein the one or more nucleotide changes are at least 7 base pairs downstream of the position of the nick site in the second strand in the double-stranded target DNA sequence, and wherein the DNA synthesis template is at least 12 bases in length.

17. The method of any one of claims 1-16, wherein the Cas9 nickase is derived from a *Streptococcus pyogenes* Cas9, a *Staphylococcus aureus* Cas9, a *Neisseria meningitis* Cas9, a *Streptococcus thermophilis* Cas9, or a *Treponema denticola* Cas9.

18. The method of any one of claims 1-17, wherein the DNA synthesis template is from 3 nucleotides to 500 nucleotides in length.

19. The method of any one of claims 1-18, wherein the DNA synthesis template is from 8 nucleotides to 31 nucleotides in length.

20. The method of any one of claims 1-19, wherein the primer binding site is from 8 nucleotides to 17 nucleotides in length.

21. The method of any one of claims 1-20, wherein the spacer sequence is about 20 nucleotides in length.