REVERSE TRANSCRIPTION-MEDIATED GENE EDITING SYSTEMS AND USES THEREOF

CROSS REFERENCE TO RELATED APPLICATIONS

This application claims the benefit of the filing dates of U.S. Provisional Application No. 63/580,188, filed September 1, 2023, U.S. Provisional Application No. 63/580,168, filed September 1, 2023, U.S. Provisional Application No. 63/553,974, filed February 15, 2024; U.S. Provisional Application No. 63/638,559, filed April 25, 2024. Each of the priority applications is incorporated by reference herein in their entities.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been filed electronically in XML format and is hereby incorporated by reference in its entirety. Said XML copy, created on August 28, 2024, is named 063586-525001WO_SeqList_ST26.xml and is 597.0 kilobytes in size.

BACKGROUND

Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated (Cas) genes, collectively known as CRISPR-Cas or CRISPR/Cas systems, are adaptive immune systems in archaea and bacteria that defend particular species against foreign genetic elements.

Reverse transcriptases (RTs) are enzymes that generate a strand of DNA that is complementary to an RNA template. The combination of reverse transcriptases and CRISPR/Cas systems has shown great potentials in genetic editing. CRISPR-guided reverse transcription allows for introduction of desired nucleotide substitutions at a genomic site.

It is therefore of great interest to develop efficient and accurate reverse transcriptase-CRISPR gene editing systems for use in disease treatment.

SUMMARY

The present disclosure describes reverse transcriptase-CRISPR-mediated gene editing systems, which successfully introduced designed nucleotide substitutions into target genetic sites. In some examples, the gene editing systems disclosed herein involve a fusion polypeptide comprising a CRISPR nuclease fragment and a reverse transcriptase (RT) fragment, and optionally one or more nuclear localization sequences (NLS) and/or peptide linkers. The

CRISPR nuclease polypeptide can be genetically engineered to possess advantageous enzymatic activities (*e.g.*, high indel activities and/or DNA cleavage activities and precise gene editing as designed). Accordingly, the gene editing systems provided herein would be expected to show superior effectiveness in inserting desired base substitutions at a genomic site of interest.

Accordingly, one example of the present disclosure features a gene editing system comprising: (a) a fusion polypeptide comprising a CRISPR nuclease polypeptide and a reverse transcriptase (RT) polypeptide, or a first nucleic acid encoding the fusion polypeptide; and (b) an RNA molecule comprising a guide RNA (gRNA) and a reverse transcription donor RNA (RT donor RNA), or a second nucleic acid encoding the RNA molecule. The CRISPR nuclease may be the reference nuclease of SEQ ID NO: 1 or a variant thereof. The gRNA comprises a scaffold sequence recognizable by the CRISPR nuclease and a spacer sequence specific to a target sequence within a genomic site of interest, the target sequence being upstream to a protospacer adjacent motif (PAM). The RT donor RNA comprises a primer binding site (PBS) and a template sequence.

In some examples, the CRISPR nuclease polypeptide comprises the amino acid sequence of SEQ ID NO: 1. In other examples, the CRISPR nuclease is a variant of SEQ ID NO: 1, the variant comprising: (i) one or more mutations in the HNH nuclease domain or in the RuvC nuclease domain of SEQ ID NO: 1 that reduce or eliminate the nuclease activity thereof; (ii) one or more arginine and/or lysine substitutions, optionally one or more arginine substitutions; or (iii) a combination of (i) and (ii).

Accordingly, in one aspect, the present disclosure refers to a gene editing system comprising: (a) a fusion polypeptide comprising a CRISPR nuclease polypeptide and a reverse transcriptase (RT) polypeptide, or a first nucleic acid encoding the fusion polypeptide; wherein the CRISPR nuclease polypeptide comprises the amino acid sequence of SEQ ID NO: 1, or is a variant of SEQ ID NO: 1, the variant comprising: (i) one or more mutations in the HNH nuclease domain or in the RuvC nuclease domain of SEQ ID NO: 1 that reduce or eliminate the nuclease activity thereof; (ii) one or more arginine and/or lysine substitutions, optionally one or more arginine substitutions; (iii) one or more mutations for reducing PAM recognition stringency; or (iv) a combination of (i), (ii), and (iii); (b) an RNA molecule comprising a guide RNA (gRNA) and a reverse transcription donor RNA (RT donor RNA), or a second nucleic acid encoding the RNA molecule; wherein the gRNA comprises a scaffold sequence recognizable by the CRISPR nuclease and a spacer sequence specific to a target sequence within a genomic site of interest, the target sequence being upstream to a protospacer adjacent motif (PAM); and wherein the RT donor RNA comprises a primer binding site (PBS) and a template sequence. In some

examples, the fusion polypeptide comprises one or more nuclear localization signal (NLS) upstream or downstream to the CRISPR nuclease polypeptide, the RT polypeptide, or both. For example, the fusion polypeptide, from N-terminus to C-terminus, comprises a first NLS, the CRISPR nuclease polypeptide, the RT polypeptide, and a second NLS. In some examples, the fusion polypeptide may further comprise a peptide linker between the CRISPR nuclease polypeptide and the RT polypeptide.

In some examples, the fusion polypeptide comprises a first peptide linker located between the CRISPR nuclease polypeptide and the RT polypeptide. Further, the fusion polypeptide may comprise a first NLS, a second NLS, which are located at the N-terminus and/or the C-terminus of the fusion polypeptide. In some examples, the fusion polypeptide may further comprise additional NLSs, for example, a third NLS and optionally a fourth NLS. Alternatively or in addition, the fusion polypeptide may further comprise a second peptide linker and optionally a third peptide linker. These peptide linkers may be located (*e.g.*, connecting) the CRISPR nuclease polypeptide and/or the RT polypeptide, and the first and/or second NLS. In

some examples, these peptide linkers may be located (*e.g.*, connecting) two NLSs.

Provided below are specific exemplary configurations for the fusion polypeptide disclosed herein (from N-terminus to C-terminus):

(i) the first NLS, the second NLS, the CRISPR nuclease polypeptide, the first peptide linker, and the RT polypeptide;

(ii) the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the first NLS, and the second NLS;

(iii) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the third NLS, the second peptide linker, and the second NLS;

(iv) the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the third NLS, the second NLS, the second peptide linker, and the first NLS;

(v) the first NLS, the second peptide linker, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the third peptide linker, and the second NLS;

(vi) the first NLS, the second peptide linker, the CRISPR nuclease, the first peptide linker, the RT polypeptide, the third peptide linker, the third NLS, the fourth NLS, and the second NLS;

(vii) the first NLS, the second peptide linker, the RT polypeptide, the first peptide linker, the CRISPR nuclease polypeptide, the third peptide linker, the third NLS, the fourth NLS, and the second NLS;

(viii) the RT polypeptide, the first peptide linker, the CRISPR nuclease polypeptide, the third NLS, the second NLS, the second peptide linker, and the first NLS;

(ix) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the second peptide linker, the second NLS, and the RT polypeptide;

(x) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the third NLS, the second peptide linker, the RT polypeptide, the third peptide linker, and the second NLS;

(xi) the first NLS, the RT polypeptide, the first peptide linker, the second NLS, the second peptide linker, and the CRISPR nuclease polypeptide,

(xii) the first NLS, the RT polypeptide, the first peptide linker, the third NLS, the second peptide linker, the CRISPR nuclease polypeptide, the third peptide linker, and the second NLS;

(xiii) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, and the second NLS; or

(xiv) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the second peptide linker, the third peptide linker, and the second NLS.

In specific examples, the fusion polypeptide may have the configuration (from N-terminus to C-terminus) of (iv), (v), (vii), (ix), or (x). See also **Table 16** below.

In some examples, wherein the peptide linker(s) between the CRISPR nuclease polypeptide and the RT polypeptide is about 20-80 amino acids in length.

In some examples, the CRISPR nuclease polypeptide comprises the variant of SEQ ID NO: 1. For example, the variant of SEQ ID NO: 1 may comprise one or more mutations in the HNH nuclease domain at positions D844, H845, and/or N868 relative to SEQ ID NO: 1. In one example, the mutation is at position H845 (*e.g.*, H845A substitution). In some examples, the mutation at D844 is an amino acid substitution of D844A, D844G, D844L, or D844S. In some examples, the mutation at H845 is an amino acid substitution of H845A, H845G, H845L, or H845S. In a specific example, the CRISPR nuclease polypeptide comprises the mutation at position H845 (e.g., H845A) relative to SEQ ID NO: 1 (*e.g.*, comprising the amino acid sequence of SEQ ID NO: 32). In some examples, the mutation at N868 is an amino acid substitution of N868A, N868G, N868L, or N868S.

Alternatively or in addition, the CRISPR nuclease polypeptide comprises a bridge helix (BH) domain, a nucleic acid recognition (REC) domain, a phosphate lock loop (PLL), a wedge (WED) domain, and a PAM-interacting (PID) domain, and wherein one or more arginine and/or lysine substitutions, optionally arginine substitutions, are located in the BH domain, in the REC domain, in the PLL domain, in the WED domain, in the PID domain, or a combination thereof. For example, the CRISPR nuclease polypeptide contains up to 20 arginine and/or lysine substitutions relative to the reference CRISPR nuclease. In specific examples, the CRISPR nuclease polypeptide contains up to 15 arginine and/or lysine substitutions relative to the reference CRISPR nuclease. In more specific examples, the one or more arginine and/or lysine substitutions are at positions K736, L784, Q812, N813, I857, and/or A919.

In some examples, the CRISPR nuclease polypeptide contains at least two arginine and/or lysine substitutions relative to the reference CRISPR nuclease. The two arginine and/or lysine substitutions are at positions K736, L784, Q812, N813, I857, and/or A919. In some examples, the CRISPR nuclease polypeptide contains arginine and/or lysine substitutions at the following positions relative to the reference CRISPR nuclease:

(a) I857, L784, and K736 (e.g., I857R, L784R, and K736R);

(b) I857, A919, and K736 (I857R, A919R, and K736R);

(c) I857, N813, and L784 (I857R, N813R, and L784R);

(d) I857, L784, and A919 (I857R, L784R, and A919R);

(e) I857, N813, and K736 (I857R, N813R, and K736R);

(f) I857 and N813 (I857R and N813R);

(g) L784, A919, and K736 (L784R, A919R, and K736R);

(h) I857 and L784 (I857R and L784R); or

(i) I857 and A919 (I857R and A919R).

In specific examples, the CRISPR nuclease polypeptide comprises the arginine substitutions of I857R, L784R, and K736R relative to SEQ ID NO: 1.

In some examples, the CRISPR nuclease polypeptide disclosed herein may further comprise one or more mutations that enhance double-strand nuclease activity relative to the reference CRISPR nuclease, in which the mutations are introduced. Examples are provided in **Table 19** below.

Alternatively or in addition, the engineered CRISPR nuclease polypeptide disclosed herein may comprise or further comprise the one or more mutations for reducing PAM recognition stringency. In some examples, the one or more mutations for reducing PAM recognition stringency may be at position D61, A68, H494, L1117, D1144, S1145, G1227, E1228, S1327, A1332, R1343, R1345, and/or T1347 of SEQ ID NO: 1. In some examples, such mutations may comprise: (i) one or more arginine and/or lysine substitutions, optionally arginine substitutions, at position D61, A68, H494, L1117, G1227, S1327, A1332, and/or T1347 of SEQ ID NO: 1; (ii) one or more amino acid substitutions at position D1144, S1145, E1228, R1343, and/or R1345, of SEQ ID NO: 1; or (iii) a combination of (i) and (ii). In specific examples, the one or more amino acid substitutions of (ii) may comprise optionally D1144L, S1145W, E1228Q, R1343P, R1345V, and/or R1345Q relative to SEQ ID NO: 1.

In specific examples, the engineered CRISPR nuclease polypeptide with less PAM recognition stringency may comprise the following combination of mutations: L1117R, D1144V, G1227R, E1228F, A1332R, R1345V, T1347R, and A68R relative to SEQ ID NO: 1. In other specific examples, the engineered CRISPR nuclease polypeptide with less PAM recognition stringency may comprise the following combination of mutations: L1117R, D1144V, G1227R, E1228F, A1332R, R1345V, T1347R, and D61R relative to SEQ ID NO: 1. In yet other specific examples, the engineered CRISPR nuclease polypeptide with less PAM recognition stringency may comprise the following combination of mutations: L1117R, D1144V, G1227R, E1228F, A1332R, R1345V, T1347R, and H494R relative to SEQ ID NO: 1. Other exemplary engineered CRISPR nuclease polypeptides can be found in **Table 20**, each

of which is within the scope of the present disclosure.

The engineered CRISPR nuclease polypeptide as disclosed herein recognize a PAM sequence of 5'-NDR-3', in which N represents A, C, G, or U, D represents A, G, or T, and R represents G or A. In some examples, the engineered CRISPR nuclease polypeptides having reduced PAM recognition stringency as disclosed herein may recognize a PAM sequence of 5'-NGN-3.' See **Example 10** below. In some examples, the PAM is 5'-NRG-3' or 5'-NRR-3', in which N and R are defined herein. In some specific examples, the PAM is 5'-NGG-3', in which N represents any nucleotide. In other specific examples, the PAM can be 5'-TGC-3' or 5'-GGA-3'.

Any of the CRISPR nuclease polypeptides may comprise the arginine and/or lysine substitutions disclosed herein, any of the nickase mutations in either the HNH or RuvC nuclease domains also disclosed herein, any of the mutations leading to reduced PAM recognition stringency, or a combination thereof. For example, the CRISPR nuclease polypeptide may comprise (a) the one or more nickase mutations in the HNH nuclease domain at positions D844, H845, and/or N868 relative to SEQ ID NO: 1 (*e.g.*, the mutation is at position H845); and (b) one or more arginine and/or lysine substitutions relative to SEQ ID NO: 1 (*e.g.*, at positions I857, L784, and K736). For example, in some examples, the CRISPR nuclease polypeptide may comprise (*e.g.*, consists of) a nickase mutation at position H845 (*e.g.*, an H845A mutation) and an arginine and/or lysine substitution at position I857 (*e.g.,* an I857R substitution) relative to SEQ ID NO: 1. In other examples, the CRISPR nuclease polypeptide may comprise or further comprise the one or more mutations that result in reduced PAM recognition stringency (e.g., at positions L1117, D1144, G1227, E1228, A1332, R1345, and/or T1347 of SEQ ID NO: 1, and optionally at one or more positions of D61, A68, and H494 of SEQ ID NO: 1).

Any of the CRISPR nuclease polypeptide provided herein may comprise an amino acid sequence at least 90% identical to SEQ ID NO: 1. In some examples, the CRISPR nuclease polypeptide comprises an amino acid sequence at least 95% identical to SEQ ID NO: 1. In yet other examples, the CRISPR nuclease polypeptide comprises an amino acid sequence at least 98% identical to SEQ ID NO: 1.

The RT polypeptide is Moloney Murine Leukemia Virus (MMLV)-RT or a variant thereof. In one example, the MMLV-RT comprises the amino acid sequence of SEQ ID NO: 53.

In some examples, the gene editing system disclosed herein comprises the fusion polypeptide. Alternatively, the system comprises the first nucleic acid encoding the fusion polypeptide. In some examples, the first nucleic acid is located on a vector, which optionally is a viral vector. In other examples, the first nucleic acid is a first messenger RNA (mRNA).

In any of the gene editing systems provided herein, the spacer sequence in the gRNA of (b) can be 15-30-nucleotide in length. In one example, the spacer sequence may be 15-20-nucleotide in length. In one specific example, the spacer sequence may be about 17-nucleotide in length.

In some examples, the scaffold sequence comprises a nucleotide sequence at least 85% identical to SEQ ID NO: 2. In one example, the scaffold sequence comprises the nucleotide sequence of SEQ ID NO: 2.

In some examples, the PBS in the RT donor RNA portion of the RNA molecule is 5-50-nucleotide in length. In some examples, the PBS is 5-20-nucleotide in length. In specific examples, the PBS can be 7-17-nucleotide in length. In some examples, the PBS binds a PBS-targeting site that is adjacent to or overlaps with the target sequence. For example, the PBS-targeting site is adjacent to or overlaps with the target sequence. In other examples, the PBS-targeting site is adjacent to the 5' of the PAM.

In some examples, the template sequence in the RT donor RNA portion of the RNA molecule can be 5-100-nucleotide in length. In some examples, the template sequence can be 15-25-nucleotide in length. In some examples, the template sequence in the RT donor RNA is homologous to the genomic site of interest and comprises one or more nucleotide variations relative to the genomic site of interest. In some examples, at least one nucleotide variation may be located within the target sequence. Alternatively, at least one nucleotide variation may be located in the PAM.

In some examples, any of the RNA molecules of (b) provided herein may further comprise a 3' end extension. In some examples, the RNA molecule may further comprise a 5' end protection fragment, a 3' protection fragment, or both, each of the 5' end protection fragment and the 3' end protection fragment forming a secondary structure, which optionally is a hairpin, a circularization, a pseudoknot, or a triplex structure.

In some examples, the RNA molecule of (b) comprises, from 5' to 3': the spacer sequence, the scaffold sequence, the template sequence, and the PBS. In other examples, the RNA molecule may comprise, from 5' to 3', the spacer sequence, the scaffold sequence, the template sequence, the PBS, and the 3' extension.

In some examples, the gene editing system disclosed herein may comprise the RNA molecule that comprises the gRNA, the RT donor RNA, and optionally one or more of the additional elements disclosed herein. Alternatively, the gene editing system may comprise the nucleic acid encoding the RNA molecule. In some examples, the nucleic acid is located on a vector, which optionally is a viral vector.

In some examples, the gene editing system disclosed herein may comprise one or more lipid nanoparticles (LNPs) associated with one or more of elements (a)-(b). Alternatively, the gene editing system may comprise one or more viral vectors, for example, optionally one or more adeno-associated viral (AAV) vectors encoding one or more of elements (a)-(b).

Also described herein in another aspect is a pharmaceutical composition comprising any of the gene editing systems provided herein. Also described herein in another aspect is a kit comprising the elements (a) and (b) of the gene editing system as disclosed herein. Also described herein in another aspect is the gene editing system as disclosed herein for use in editing a target gene in a subject.

In other examples, the present disclosure features a gene editing method, comprising delivering the gene editing system disclosed herein to a host cell to edit a genomic site targeted by the gRNA of the gene editing system. In some examples, the host cell is cultured *in vitro*. In other examples, the host cell is located in a subject who needs the gene editing. In one aspect, the present disclosure refers to an *in vitro* gene editing method, comprising delivering the gene editing system as disclosed herein to a host cell cultured *in vitro* to edit a genomic site targeted by the gRNA of the gene editing system

Further, in another aspect, the present disclosure refers to a fusion polypeptide, comprising any of the CRISPR nuclease polypeptide set forth herein and any of the reverse transcriptase polypeptide also set forth herein. Such a fusion polypeptide may comprise the amino acid sequence of SEQ ID NO: 55 or 57.

In addition, in another aspect, the present disclosure refers to a nucleic acid encoding the fusion polypeptide disclosed herein. Such a nucleic acid may comprise the nucleotide sequence of SEQ ID NO: 54 or 56. In some examples, the nucleic acid is a vector, such as an expression vector.

The details of one or more examples of the invention are set forth in the description below. Other features or advantages of the present invention will be apparent from the following drawings and detailed description of several examples, and also from the appended claims.

BRIEF DESCRIPTION OF THE DRAWINGS

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 the drawing in combination with the detailed description of specific embodiments

presented herein.

FIG. 1 is a diagram showing gene editing efficacy of reference CRISPR nuclease SEQ ID NO: 1 on exemplary target genes AAVS1, EMX1, and VEGFA.

FIGs. 2A-2D include gel images showing quantification of nuclease activities. **FIG. 2A**: a gel image captured using a 700 nm channel showing *in vitro* cleavage of the target strand (labelled on the 5' end with an IR700 dye) of the target DNA substrate by the reference CRISPR nuclease, putative HNH-knockout nickases, or putative RuvC-knockout nickases. **FIG. 2B**: a gel image captured using an 800 nm channel showing *in vitro* cleavage of the non-target strand

WHAT IS CLAIMED IS:

1. A gene editing system comprising:

(a) a fusion polypeptide comprising a CRISPR nuclease polypeptide and a reverse transcriptase (RT) polypeptide, or a first nucleic acid encoding the fusion polypeptide; wherein the CRISPR nuclease polypeptide comprises the amino acid sequence of SEQ ID NO: 1, or is a variant of SEQ ID NO: 1, the variant comprising:

(i) one or more mutations in the HNH nuclease domain or in the RuvC nuclease domain of SEQ ID NO: 1 that reduce or eliminate the nuclease activity thereof;

(ii) one or more arginine and/or lysine substitutions, optionally one or more arginine substitutions;

(iii) one or more mutations for reducing PAM recognition stringency; or

(iv) a combination of any of (i), (ii), and (iii);

(b) an RNA molecule comprising a guide RNA (gRNA) and a reverse transcription donor RNA (RT donor RNA), or a second nucleic acid encoding the RNA molecule;

wherein the gRNA comprises a scaffold sequence recognizable by the CRISPR nuclease and a spacer sequence specific to a target sequence within a genomic site of interest, the target sequence being upstream to a protospacer adjacent motif (PAM); and

wherein the RT donor RNA comprises a primer binding site (PBS) and a template sequence.

2. The gene editing system of claim 1, wherein the fusion polypeptide further comprises one or more nuclear localization signals (NLSs) upstream or downstream to the CRISPR nuclease polypeptide, the RT polypeptide, or both.

3. The gene editing system of claim 2, wherein the fusion polypeptide, from N-terminus to C-terminus, comprises a first NLS, the CRISPR nuclease polypeptide, the RT polypeptide, and a second NLS; optionally wherein the fusion polypeptide further comprises a peptide linker between the CRISPR nuclease polypeptide and the RT polypeptide; or

wherein the fusion polypeptide comprises a first peptide linker located between the CRISPR nuclease polypeptide and the RT polypeptide, and wherein the fusion polypeptide comprises a first NLS, a second NLS, which are located at the N-terminus and/or the C-

terminus of the fusion polypeptide.

4. The gene editing system of claim 3, wherein the fusion polypeptide further comprises a third NLS and optionally a fourth NLS; or

wherein the fusion polypeptide further comprises a second peptide linker connecting the CRISPR nuclease polypeptide or the RT polypeptide, and the first or second NLS(s); optionally wherein the fusion polypeptide further comprises a third peptide linker connecting two NLSs.

5. The gene editing system of any one of claims 3-4, wherein the fusion polypeptide comprises, from N-terminus to C-terminus,

(i) the first NLS, the second NLS, the CRISPR nuclease polypeptide, the first peptide linker, and the RT polypeptide;

(ii) the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the first NLS, and the second NLS;

(iii) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the third NLS, the second peptide linker, and the second NLS;

(iv) the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the third NLS, the second NLS, the second peptide linker, and the first NLS;

(v) the first NLS, the second peptide linker, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the third peptide linker, and the second NLS;

(vi) the first NLS, the second peptide linker, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the third peptide linker, the third NLS, the fourth NLS, and the second NLS;

(vii) the first NLS, the second peptide linker, the RT polypeptide, the first peptide linker, the CRISPR nuclease polypeptide, the third peptide linker, the third NLS, the fourth NLS, and the second NLS;

(viii) the RT polypeptide, the first peptide linker, the CRISPR nuclease polypeptide, the third NLS, the second NLS, the second peptide linker, and the first NLS;

(ix) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the second peptide linker, the second NLS, and the RT polypeptide;

(x) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the third NLS, the second peptide linker, the RT polypeptide, the third peptide

linker, and the second NLS;

(xi) the first NLS, the RT polypeptide, the first peptide linker, the second NLS, the second peptide linker, and the CRISPR nuclease polypeptide,

(xii) the first NLS, the RT polypeptide, the first peptide linker, the third NLS, the second peptide linker, the CRISPR nuclease polypeptide, the third peptide linker, and the second NLS;

(xiii) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, and the second NLS; or

(xiv) the first NLS, the CRISPR nuclease polypeptide, the first peptide linker, the RT polypeptide, the second peptide linker, the third peptide linker, and the second NLS.

6. The gene editing system of any one of claims 3-5, wherein the peptide linker(s) between the CRISPR nuclease polypeptide and the RT polypeptide is about 20-80 amino acids in length.

7. The gene editing system of any one of claims 1-6, wherein the CRISPR nuclease polypeptide is the variant of SEQ ID NO: 1, which comprises one or more mutations in the HNH nuclease domain at positions D844, H845, and/or N868 relative to SEQ ID NO: 1; optionally wherein the mutation is at position H845, which optionally is H845A; and/or

wherein the CRISPR nuclease polypeptide comprises a bridge helix (BH) domain, a nucleic acid recognition (REC) domain, a phosphate lock loop (PLL), a wedge (WED) domain, and a PAM-interacting (PID) domain; and wherein one or more arginine and/or lysine substitutions are located in the BH domain, in the REC domain, in the PLL domain, in the WED domain, in the PID domain, or a combination thereof.

8. The gene editing system of claim 7, wherein the CRISPR nuclease polypeptide comprises the one or more mutations in the HNH nuclease domain, and wherein:

(a) the mutation at D844 is an amino acid substitution of D844A, D844G, D844L, or D844S;

(b) the mutation at H845 is an amino acid substitution of H845A, H845G, H845L, or H845S; and

(c) the mutation at N868 is an amino acid substitution of N868A, N868G, N868L, or N868S; and/or

wherein the CRISPR nuclease polypeptide contains up to 20 arginine and/or lysine substitutions relative to SEQ ID NO: 1; optionally wherein the CRISPR nuclease polypeptide contains up to 15 arginine and/or lysine substitutions relative to SEQ ID NO: 1; preferably wherein the one or more arginine and/or lysine substitutions are at positions K736, L784, Q812, N813, I857, and/or A919 of SEQ ID NO: 1, which optionally is at position I857, optionally I857R.

9. The gene editing system of 7, wherein the CRISPR nuclease variant is any one of those listed in **Table 19** or **Table 20**.

10. The gene editing system of claim 8, wherein the CRISPR nuclease polypeptide contains at least two arginine and/or lysine substitutions relative to SEQ ID NO: 1, and wherein the at least two arginine and/or lysine substitutions are at positions K736, L784, Q812, N813, I857, and/or A919 of SEQ ID NO: 1.

11. The gene editing system of claim 10, wherein the CRISPR nuclease polypeptide contains arginine and/or lysine substitutions at the following positions relative to SEQ ID NO: 1:

(a) I857, L784, and K736;
(b) I857, A919, and K736;
(c) I857, N813, and L784;
(d) I857, L784, and A919;
(e) I857, N813, and K736;
(f) I857 and N813;
(g) L784, A919, and K736;
(h) I857 and L784; and
(i) I857 and A919;

optionally wherein the CRISPR nuclease polypeptide comprises the following arginine substitutions relative to SEQ ID NO: 1:

(a) I857R, L784R, and K736R;
(b) I857R, A919R, and K736R;
(c) I857R, N813R, and L784R;
(d) I857R, L784R, and A919R;
(e) I857R, N813R, and K736R;

(f) I857R and N813R;

(g) L784R, A919R, and K736R;

(h) I857R and L784R; and

(i) I857R and A919R;

optionally wherein the CRISPR nuclease polypeptide comprises the arginine substitutions of (a).

12. The gene editing system of claim 1, wherein the CRISPR nuclease polypeptide comprises a nickase mutation at position H845, optionally H845A, and an arginine and/or lysine substitution at position I857, optionally I857R relative to SEQ ID NO: 1.

13. The gene editing system of any one of claims 1-12, wherein the CRISPR nuclease polypeptide comprises the one or more mutations for reducing PAM recognition stringency, optionally wherein the one or more mutations are at positions D61, A68, H494, L1117, D1144, S1145, G1227, E1228, S1327, A1332, R1343, R1345, and/or T1347 of SEQ ID NO: 1.

14. The gene editing system of any one of claims 7-13, wherein the CRISPR nuclease polypeptide comprises (a) the one or more mutations in the HNH nuclease domain at positions D844, H845, and/or N868 relative to SEQ ID NO: 1; optionally wherein the mutation is at position H845; and (b) one or more arginine and/or lysine substitutions relative to SEQ ID NO: 1, optionally wherein the arginine and/or lysine substitutions are at positions I857, L784, and K736.

15. The gene editing system of any one of claims 7-14, wherein the CRISPR nuclease polypeptide comprises an amino acid sequence at least 90% identical to SEQ ID NO: 1.

16. The gene editing system of any one of claims 1-15, wherein the RT polypeptide is Moloney Murine Leukemia Virus (MMLV)-RT, optionally the MMLV-RT comprises the amino acid sequence of SEQ ID NO: 53.

17. The gene editing system of claim 1, wherein the fusion polypeptide is set forth in **Table 8** or **Table 17**.

18. The gene editing system of any one of claims 1-17, wherein the system comprises the fusion polypeptide.

19. The gene editing system of any one of claims 1-17, wherein the system comprises the first nucleic acid encoding the fusion polypeptide.

20. The gene editing system of claim 19, wherein the first nucleic acid is located on a vector, which optionally is a viral vector; or

wherein the first nucleic acid is a first messenger RNA (mRNA).

21. The gene editing system of any one of claims 1-20, wherein the spacer sequence in the gRNA of (b) is 15-30-nucleotide in length; optionally 15-20-nucleotide in length;

wherein the PAM is 5'-NDR-3' or 5'-NGN-3', in which N represents any nucleotide, D represents A, G, or T, and R represents G or A; optionally wherein the PAM is 5'-NRG-3' or 5'-NRR'3', in which N and Rare defined herein; preferably wherein the PAM is 5'-NGG-3', in which N represents any nucleotide;

wherein the scaffold sequence comprises a nucleotide sequence at least 85% identical to SEQ ID NO: 2, optionally wherein the scaffold sequence comprises the nucleotide sequence of SEQ ID NO: 2; and/or

wherein the PBS in the RT donor RNA is 5-50-nucleotide in length; optionally 5-20-nucleotide in length; optionally wherein the PBS binds a PBS-targeting site that is adjacent to or overlaps with the target sequence, or wherein the PBS-targeting site is adjacent to the 5' of the PAM, optionally wherein the 3' end nucleotide of the PBS-targeting site is about 2-15 nucleotides upstream to the PAM.

22. The gene editing system of any one of claims 1-21, wherein the template sequence in the RT donor RNA is 5-100-nucleotide in length; optionally 15-25-nucleotide in length; and/or

wherein the template sequence in the RT donor RNA is homologous to the genomic site of interest and comprises one or more nucleotide variations relative to the genomic site of interest;

optionally wherein at least one nucleotide variation is located within the target sequence; and/or wherein at least one nucleotide variation is located in the PAM.

23. The gene editing system of any one of claims 1-22, wherein the RNA molecule of (b) further comprises a 3' end extension; and/or

wherein the RNA molecule of (b) further comprises a 5' end protection fragment, a 3' protection fragment, or both, each of the 5' end protection fragment and the 3' end protection fragment forming a secondary structure, which optionally is a hairpin, a pseudoknot, a circularization, or a triplex structure.

24. The gene editing system of any one of claims 1-23, wherein the RNA molecule of (b) comprises, from 5' to 3':

(i) the spacer sequence, the scaffold sequence, the template sequence, and the PBS; or

(ii) the spacer sequence, the scaffold sequence, the template sequence, the PBS, and the 3' end extension.

25. The gene editing system of any one of claims 1-24, wherein the system comprises the RNA molecule of (b).

26. The gene editing system of any one of claims 1-24, wherein the system comprises the second nucleic acid encoding the RNA molecule;

optionally wherein the second nucleic acid is located on a vector, which optionally is a viral vector.

27. The gene editing system of any one of claims 1-26, wherein the system comprises one or more lipid nanoparticles (LNPs) associated with one or more of elements (a)-(b); or

wherein the system comprises one or more viral vectors, optionally one or more adeno-associated viral (AAV) vectors encoding one or more of elements (a)-(b).

28. A pharmaceutical composition comprising the gene editing system of any one of claims 1-27.

29. A kit comprising the elements (a)-(b) of the gene editing system set forth in any one of claims 1-27.

30. An *in vitro* gene editing method, comprising delivering the gene editing system of any one of claims 1-27 to a host cell cultured *in vitro* to edit a genomic site targeted by the gRNA of the gene editing system.

31. The gene editing system of any one of claims 1-27 for use in editing a target gene in a subject.

32. A fusion polypeptide, comprising a CRISPR nuclease polypeptide set forth in any one of claims 1-15 and a reverse transcriptase polypeptide set forth in claim 1 or claim 16.

33. The fusion polypeptide of claim 32, which comprises the amino acid sequence of SEQ ID NO: 55 or 57.

34. A nucleic acid encoding the fusion polypeptide of claim 32 or claim 33.

35. The nucleic acid of claim 34, which is a vector, optionally an expression vector.