COMPOSITIONS AND METHODS FOR EPIGENETIC EDITING

CROSS REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 63/354,931, filed June 23, 2022, which is incorporated by reference in its entirety.

BACKGROUND

[0002] Genome editing has been considered a promising therapeutic approach for treatment of genetic disease for over a decade. However, manipulation on the DNA level remains risky given the potential for undesired double stranded breaks, heterogenous repair including large and small insertions and deletions at the intended site, and toxicity.

SUMMARY

[0002a] In one aspect, provided herein is an epigenetic editing system comprising: (a) a fusion protein comprising: a DNA-binding domain, and a DNMT3L domain, wherein the DNMT3L domain is from a species selected from the group consisting of *Ailuropoda melanoleuca, Carlito syrichta, Meriones unguiculatus, Ochotona princeps, Neosciurus carolinensis, Bison bison, Equus przewalskii, Mus caroli and Pan troglodytes*; or (b) a nucleic acid molecule encoding the fusion protein of (a).

[0003] Provided herein are compositions for epigenetic modification related to epigenetic editing systems and methods of using the same to generate epigenetic modification in target genomes, including those in host cells and organisms, without introducing changes to genomic sequences.

[0004] Described herein is an epigenetic editing system comprising: (a) a fusion protein comprising a DNA-binding domain, a DNA methyltransferase (DNMT) domain, a repressor domain and two nuclear localization sequences (NLSs), wherein each of the two NLSs is positioned at the amino (N) terminus or at the carboxy (C) terminus of the fusion protein; or (b) a nucleic acid molecule encoding the fusion protein of (a).

[0005] In some examples, the fusion protein comprises one or more NLSs at the C terminus and one or more NLSs at the N terminus of the fusion protein, optionally wherein the fusion protein comprises two NLSs at the N terminus and two NLSs at the C terminus of the fusion protein. In some examples, the DNMT domain is from a bacterial species. In some examples, the DNMT domain is a mammalian DNMT domain. In some examples, the DNMT domain is a mouse DNMT domain. In some examples, the DNMT domain is a human DNMT domain. In some examples, the DNMT domain is a DNMT3A domain. In some

examples, the DNMT domain is a DNMT3L domain. In some examples, the fusion protein comprises both a DNMT3A domain and a DNMT3L domain.

[0006] In some examples, the DNMT3L domain is from a species selected from the group consisting of *Ailuropoda melanoleuca*, *Carlito syrichta*, *Meriones unguiculatus*, *Ochotona princeps*, *Neosciurus carolinensis*, *Bison bison*, *Equus przewalskii*, *Mus caroli* and *Pan troglodytes*; optionally comprising one of SEQ ID NOs: 72-80 and 101-109, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto.

[0007] In some examples, the repressor domain comprises a KRAB domain of a ZFP28, ZN627, KAP1, MeCP2, HP1b, CBX8, CDYL2, TOX, Tox3, Tox4, EED, RBBP4, RCOR1, or SCML2 protein, or a fusion of the N- and C-terminal regions of ZIM3 and KOX1 KRAB domains.

[0008] Described herein is an epigenetic editing system comprising: (a) a fusion protein comprising a DNA-binding domain, a DNA methyltransferase (DNMT) domain from a bacterial species, wherein the DNMT domain is fused to the N terminus of the DNA binding domain; or (b) a nucleic acid molecule encoding the fusion protein of (a).

[0009] In some examples, the fusion protein functions to methylate mammalian target DNA in a cell. In some examples, the DNMT domain of the fusion protein does not comprise any one of SEQ ID NOs: 81-93. In some examples, the bacterial species is not *M. penetrans*, *S. monbiae*, *H. parainfluenzae*, *A. luteus*, *H. aegyptius*, *H. haemolyticus*, *Moraxella*, *E. coli*, *T. aquaticus*, *C. crescentus*, or *C. difficile*.

[0010] In some examples, the DNMT domain from the bacterial species is derived from (a) M.Sss1, optionally comprising SEQ ID NO: 40, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto; NQZ29229, optionally comprising SEQ ID NO: 41, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto; WP_131599610, optionally comprising SEQ ID NO: 42, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto; or WP_208057179, optionally comprising SEQ ID NO: 43, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto.

[0011] Described herein is an epigenetic editing system comprising: (a) a fusion protein comprising a DNA-binding domain, and a DNMT3L domain, wherein the DNMT3L domain is from a species selected from the group consisting of *Ailuropoda melanoleuca*, *Carlito syrichta*, *Meriones unguiculatus*, *Ochotona princeps*, *Neosciurus carolinensis*, *Bison bison*, *Equus przewalskii*, *Mus caroli* and *Pan* troglodytes; optionally comprising one of SEQ ID NOs: 72-80 and 101-109, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto; or (b) a nucleic acid molecule encoding the fusion protein of (a).

[0012] In some examples, (i) the fusion protein further comprises a repressor domain, or (ii) the system further comprises an additional fusion protein comprising a DNA-binding domain and a repressor domain, or a nucleic acid molecule encoding the additional fusion protein. In some examples, the repressor domain comprises a KRAB domain, optionally derived from KOX1, ZIM3, ZFP28, or ZN627.

[0013] In some examples, the KRAB domain is derived from human KOX1, optionally comprising SEQ ID NO: 94, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto, or comprising SEQ ID NO: 100, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto. In some examples, the repressor domain is derived from KAP1, MECP2, HP1a/CBX5, HP1b, CBX8, CDYL2, TOX, TOX3, TOX4, EED, EZH2, RBBP4, RCOR1, or SCML2.

[0014] Described herein is an epigenetic editing system comprising: (a) a fusion protein comprising a DNA-binding domain, and a repressor domain, wherein the repressor domain comprises a KRAB domain of a ZFP28, ZN627, KAP1, MeCP2, HP1b, CBX8, CDYL2, TOX, Tox3, Tox4, EED, RBBP4, RCOR1, or SCML2 protein, or a fusion of the N- and C-terminal regions of ZIM3 and KOX1 KRAB domains; or (b) a nucleic acid molecule encoding the fusion protein of (a).

[0015] In some examples, the repressor domain comprises one of SEQ ID NOs: 44-57, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto. In some examples, the fusion protein further comprises a DNA methyltransferase (DNMT) domain, or the system further comprises an additional fusion protein comprising a DNA-binding domain and a DNMT domain, or a nucleic acid molecule encoding the additional fusion protein.

[0016] In some examples, the fusion protein or the system comprises a human DNMT3A domain and a human DNMT3L domain, or a human DNMT3A domain and a mouse DNMT3L domain, optionally (a) wherein the human DNMT3A domain comprises SEQ ID NO: 12, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto; (b) wherein the human DNMT3L domain comprises SEQ ID NO: 13, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto, and/or (c) wherein the mouse DNMT3L domain comprises SEQ ID NO: 15, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto.

[0017] In some examples, the epigenetic editing system further comprises a fusion protein comprising, from N terminus to C terminus, one or more NLSs, the DNMT3A domain, the DNMT3L domain, the DNA-binding domain, the repressor domain, and one or more NLSs, optionally wherein the fusion protein comprises a peptide linker between adjacent domains. In some examples, the fusion protein comprises, from N terminus to C terminus, two NLSs, the DNMT3A domain, ADD, the DNMT3L domain, a first peptide linker, the DNA-binding domain, a second peptide linker, the repressor domain, and two NLSs.

[0018] In some examples, the DNMT3A domain is a human DNMT3A domain and/or the DNMT3L domain is a human DNMT3L domain. In some examples, the repressor domain is a KRAB domain from mammalian, optionally human, ZFP28, ZF627, or KOX1.

[0019] In some examples, the first peptide linker is XTEN80 (SEQ ID NO: 3) and/or the second peptide linker is XTEN16 (SEQ ID NO: 2). In some examples, the system comprises an expression construct encoding the fusion protein, wherein the expression construct comprises a WPRE sequence in a 3' noncoding region and upstream from a poly-adenylation site.

[0020] In some examples, the DNA-binding domain is a dCas9 domain. In some examples, the dCas9 domain comprises SEQ ID NO: 9, or an amino acid sequence at least 90%, optionally at least 95%, homologous thereto. In some examples, the epigenetic editing system further comprises one or more guide RNAs (gRNAs) or nucleic acid molecule(s) coding for the gRNAs. In some examples, the DNA-binding domain is a zinc finger protein (ZFP) domain.

[0021] Described herein is a method of modifying an epigenetic state of a target gene in a mammalian cell, comprising contacting the cell with the epigenetic editing system of the disclosure. Described herein is a method of modulating expression of a target gene in a mammalian cell, comprising contacting the cell with the epigenetic editing system of the disclosure. Described herein is a method of treating a disease in a subject in need thereof, comprising administering to the subject the epigenetic editing system of the disclosure.

[0022] Additional examples and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative examples of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different examples, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

[0023] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.