

72(*C9orf72*) gene target nucleic acid sequences. The Class 2 Type V CRISPR proteins and guide nucleic acids are modified for passive entry into target cells. The Class 2 Type V CRISPR proteins and guide nucleic acids are useful in a variety of methods for target nucleic acid modification of *C9orf72*-related diseases, which methods are also provided.

[0007] In one example, the present disclosure relates to CasX:guide nucleic acid systems (CasX:gNA system) and methods used to alter a target nucleic acid comprising the *C9orf72* gene with one or more mutations or that comprises the hexanucleotide repeat sequence expansion (HRS) in cells. In some examples of the disclosure, the CasX:gNA system has utility in knocking-down or knocking-out a *C9orf72* gene with one or more mutations or that comprises the hexanucleotide repeat sequence expansion (HRS), in order to reduce or eliminate expression of the *C9orf72* gene product, accumulation of the RNA from the HRS, and/or the DPR in subjects having a *C9orf72*-related disease. In other examples, the CasX:gNA system has utility in correcting the *C9orf72* gene comprising the HRS.

[0007a] In one aspect, the present disclosure refers to a system comprising a CasX variant protein and a guide nucleic acid (gNA), wherein the CasX variant protein comprises a sequence having at least 70% sequence identity to SEQ ID NO: 138 and exhibits one or more improved characteristics relative to a CasX reference protein of any one of SEQ ID NOS: 1-3, wherein the CasX variant protein is capable of forming a ribonuclear protein complex (RNP) with the gNA, and wherein the gNA comprises a targeting sequence complementary to a target nucleic acid sequence comprising a *chromosome 9 open reading frame 72 (C9orf72)* gene.

[0008] In some examples of the system, the gNA is a gRNA, or a gDNA, or a chimera of RNA and DNA, and may be a single-molecule gNA or a dual-molecule gNA. In other examples, the CasX:gNA system gNA has a targeting sequence complementary to a target nucleic acid sequence comprising a region within the *C9orf72* gene. In some examples, the targeting sequence of the gNA is selected from the group consisting of SEQ ID NOS: 309-343, 363-2100, 2295-2185, or a sequence having at least about 65%, at least about 75%, at least about 85%, or at least about 95% identity thereto. The gNA can comprise a targeting sequence comprising 14 to 30 consecutive nucleotides. In some examples, the targeting sequence of the gNA consists of 21 nucleotides. In other examples, the targeting sequence of the gNA consists of 20 nucleotides. In other examples, the targeting sequence consists of 19 nucleotides, the targeting sequence of the gNA having a sequence selected from the group consisting of SEQ ID NOS: 309-343, 363-2100, 2295-2185 with a single nucleotide removed from the 3' end of the sequence. In other examples, the targeting sequence consists of 18 nucleotides, having a sequence selected from the group consisting of SEQ ID NOS: 309-343, 363-2100, 2295-2185 with two nucleotides removed from the 3' end of the sequence. In other examples, the targeting sequence consists of 17 nucleotides, having a sequence selected from the group consisting of SEQ ID NOS: 309-343, 363-2100, 2295-2185 with three nucleotides removed from the 3' end of the sequence. In other examples, the targeting sequence consists of 16 nucleotides, having a sequence selected from the group consisting of

SEQ ID NOS: 309-343, 363-2100, 2295-2185 with four nucleotides removed from the 3' end of the sequence. In other examples, the targeting sequence consists of 15 nucleotides, having a sequence selected from the group consisting of SEQ ID NOS: 309-343, 363-2100, 2295-2185 with five nucleotides removed from the 3' end of the sequence.

[0009] In some examples of the system, the gNA has a scaffold comprising a sequence selected from the group consisting of SEQ ID NOS: 4-16 and 2101-2294, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto.

[0010] In some examples of the system, the Class 2, Type V CRISPR protein comprises a reference CasX protein having a sequence of any one of SEQ ID NOS: 1-3, a CasX variant protein having a sequence selected from the group consisting of SEQ ID NOS: 49-150, 233-235, 238-239, 240-242, and 272-281, or a sequence having at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, or at least about 99% sequence identity thereto. In these examples, a CasX variant exhibits one or more improved characteristics relative to the reference CasX protein. In some examples, the CasX protein has binding affinity for a protospacer adjacent motif (PAM) sequence selected from the group consisting of TTC, ATC, GTC, and CTC. In some examples, the CasX protein has binding affinity for the PAM sequence that is at least 1.5-fold greater compared to the binding affinity of any one of the CasX proteins of SEQ ID NOS: 1-3 for the PAM sequences selected from the group consisting of TTC, ATC, GTC, and CTC.

[0011] In some examples of the system, the CasX molecule and the gNA molecule are associated together in a ribonuclear protein complex (RNP). In a particular example, the RNP comprising the CasX variant and the gNA variant exhibits greater editing efficiency and/or binding of a target sequence in the target DNA when any one of the PAM sequences TTC, ATC, GTC, or CTC is located 1 nucleotide 5' to the non-target strand sequence having identity with the targeting sequence of the gNA in a cellular assay system compared to the editing efficiency and/or binding of an RNP comprising a reference CasX protein and a reference gNA in a comparable assay system.

[0012] In some examples, the system further comprises a donor template comprising a nucleic acid comprising at least a portion of a *C9orf72* gene, wherein the *C9orf72* gene portion

is selected from the group consisting of a *C9orf72* exon, a *C9orf72* intron, a *C9orf72* intron-exon junction, a *C9orf72* regulatory element, or combinations thereof, wherein the donor template is used to knock down or knock out the *C9orf72* gene or is used to correct the mutation in the *C9orf72* gene. In some examples, the donor template comprises a hexanucleotide repeat of a GGGGCC sequence wherein the number of repeats ranges from 10 to about 30 repeats and is utilized to replace the hexanucleotide repeat expansion of the mutant *C9orf72* gene. In some cases, the donor sequence is a single-stranded DNA template or a single stranded RNA template. In other cases, the donor template is a double-stranded DNA template.

[0013] In other examples, the disclosure relates to nucleic acids encoding the systems of any of the examples described herein, as well as vectors comprising the nucleic acids. In some examples, the vector is selected from the group consisting of a retroviral vector, a lentiviral vector, an adenoviral vector, an adeno-associated viral (AAV) vector, a herpes simplex virus (HSV) vector, a plasmid, a minicircle, a nanoplasmid, and an RNA vector. In other examples, the vector is a virus-like particle (VLP) comprising an RNP of a CasX and gNA of any of the examples described herein and, optionally, a donor template nucleic acid and a targeting moiety such as a viral-derived glycoprotein.

[0013a] In one aspect, the present disclosure refers to a nucleic acid comprising a sequence encoding: (a) a gNA comprising a targeting sequence complementary to a target nucleic acid sequence comprising a chromosome 9 open reading frame 72 (*C9orf72*) gene; and (b) a CasX variant protein comprising a sequence having at least 70% sequence identity to SEQ ID NO: 138 and exhibits one or more improved characteristics relative to a CasX reference protein of any one of SEQ ID NOS: 1-3, wherein the CasX variant protein is capable of forming a ribonuclear protein complex (RNP) with the gRNA, and wherein the CasX variant protein further comprises one or more nuclear localization signals (NLS).

[0013b] In one aspect, the present disclosure refers to a vector comprising the nucleic acid disclosed herein, wherein the vector is selected from the group consisting of a retroviral vector, a lentiviral vector, an adenoviral vector, an adeno-associated viral (AAV) vector, a herpes simplex virus (HSV) vector, a virus-like particle (VLP), a plasmid, a minicircle, a nanoplasmid, and an RNA vector.

[0013c] In one aspect, the present disclosure refers to a host cell comprising the vector disclosed herein, wherein the host cell is selected from the group consisting of BHK, HEK293, HEK293T, NS0, SP2/0, YO myeloma cells, P3X63 mouse myeloma cells, PER, PER.C6, NIH3T3, COS, HeLa, CHO, and yeast cells.

[0014] In other examples, the disclosure provides a method of modifying a *C9orf72* target nucleic acid sequence of cells of a population, wherein said method comprises introducing into the cell: a) CasX:gNA system of any of the examples disclosed herein; b) the nucleic acid of any of the examples disclosed herein; c) the vector of any of the examples disclosed herein; d) the VLP of any of the examples disclosed herein; or e) a combination of the foregoing, wherein the *C9orf72* gene target nucleic acid sequence of the cells targeted by the first gNA is modified by the CasX protein, introducing a single- or double-stranded break in the target nucleic acid sequence. In some examples of the method, the method further comprises a second gNA or a nucleic acid encoding the second gNA, wherein the second gNA has a targeting sequence complementary to a different portion of the target nucleic acid sequence. In some examples of the method, the modifying comprises introducing an insertion, deletion, substitution, duplication, or inversion of one or more nucleotides in the target nucleic acid sequence as compared to the wild-type sequence. In some cases, the method further comprises contacting the target nucleic acid with a donor template nucleic acid of any of the examples disclosed herein. In some examples, the target *C9orf72* gene for modification comprises more than 30, more than 100, more than 500, more than 700, more than 1000, or more than 1600

copies of a hexanucleotide repeat sequence GGGGCC. In some examples of the method, the donor template comprises a nucleic acid comprising at least a portion of a *C9orf72* gene for correcting (by knocking in) the mutation of the *C9orf72* gene, or comprises a sequence comprising a mutation or heterologous sequence for knocking-down or knocking out the mutant *C9orf72* such that expression of the HRS or the DPR by the cells of the population is reduced by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% in comparison to a cell where the that has not been modified. In some cases, the modifying of the target nucleic acid sequence occurs *in vivo*. In some examples, the cell is a eukaryotic cell selected from the group consisting of a rodent cell, a mouse cell, a rat cell, a primate cell, a and a non-human primate cell. In some examples, the cell is a human cell. In some examples, the cell is a selected from the group consisting of a Purkinje cell, frontal cortex neuron, motor cortex neuron, hippocampus neuron, cerebellum neuron, upper motor neuron, spinal cord neuron, spinal cord motor neuron, glial cell, and astrocytes.

[0014a] In one aspect, the present disclosure refers to a method of modifying a *chromosome 9 open reading frame 72 (C9orf72)* gene target nucleic acid sequence in a population of cells, the method comprising the steps of, introducing into the population of cells: a. the system disclosed herein; b. the nucleic acid disclosed herein; c. the vector disclosed herein; or d. a combination thereof; and modifying the *C9orf72* gene target nucleic acid sequence in the population of cells targeted by the gNA and the CasX variant protein, wherein the modifying comprises: (i) introducing a double-stranded break in the *C9orf72* target nucleic acid sequence; and/or (ii) introducing an insertion, deletion, substitution, duplication, or inversion of one or more nucleotides in the *C9orf72* target nucleic acid sequence.

[0014b] In one aspect, the present disclosure refers to a population of cells modified by: a. the system disclosed herein; b. the nucleic acid disclosed herein; c. the vector disclosed herein; or d. a combination thereof, wherein the population of cells once modified include at least one of the following characteristics relative to a population of unmodified cells: (i) at least 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90% do not express a detectable level of dipeptide repeat protein (DPR), (ii) expression of a functional *C9orf72* protein is increased by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, or at least about 90 %, and/or (iii) mutation of the *C9orf72* gene is corrected, resulting in expression of a functional *C9orf72* protein, wherein cells of the population of cells are selected from the group consisting of Purkinje cell, frontal cortex neuron, motor cortex neuron, hippocampus neuron, cerebellum neuron, upper motor neuron, spinal cord neuron, spinal cord motor neuron, glial cell, and astrocytes.

[0015] In other examples, the disclosure provides methods of modifying a *C9orf72* target nucleic acid in a population of cells of a subject wherein the target cells are contacted using vectors encoding the CasX protein and one or more gNAs comprising a targeting sequence complementary to the *C9orf72* gene, and optionally further comprising a donor template. In some cases, the vector is an Adeno-Associated Viral (AAV) vector selected from AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, AAV 44.9, AAV-Rh74, or AAVRh10. In other cases, the vector is a lentiviral vector. In other examples, the disclosure provides methods wherein the target cells are contacted using a vector wherein the vector is a virus-like particle (VLP) comprising an RNP of a CasX and gNA of any of the examples described herein and, optionally, a donor template nucleic acid. In some examples of the method, the vector is administered to a subject at a therapeutically effective dose. The subject can be a mouse, rat, pig, non-human primate, or a human. The dose can be administered by a route of administration selected from the group consisting of subcutaneous, intradermal, intraneural, intranodal, intramedullary, intramuscular, intralumbar, intrathecal, subarachnoid, intraventricular, intracapsular, intravenous, intralymphatic, or intraperitoneal routes, wherein the administering method is injection, transfusion, or implantation.

[0016] In other examples, the disclosure provides a method of treating a *C9orf72*-related disorder in a subject, comprising modifying a gene encoding *C9orf72* gene in cells of the subject, the modifying comprising contacting said cells with: a) CasX:gNA system of any of the examples disclosed herein; b) the nucleic acid of any of the examples disclosed herein; c) the vector of any of the examples disclosed herein; d) the VLP of any of the examples disclosed herein; or e) a combination of the foregoing wherein the *C9orf72* gene of the cells targeted by the first gNA is modified by the CasX protein. In some examples, the subject is selected from the group consisting of mouse, rat, pig, non-human primate, and human. In some examples, the *C9orf72*-related disorder is ALS or FTD. In some cases, the methods of treating a subject with a *C9orf72*-related disease result in improvement in at least one clinically-relevant parameter. In other cases, the methods of treating a subject with a *C9orf72*-related disease result in improvement in at least two clinically-relevant parameters.

[0017] In other examples, the disclosure provides compositions for use in a method of treating a *C9orf72*-related disorder in a subject. In some examples, the method comprises modifying a gene encoding *C9orf72* gene in cells of the subject, the modifying comprising contacting said cells with compositions selected from: a) CasX:gNA system of any of the examples disclosed herein; b) the nucleic acid of any of the examples disclosed herein; c) the vector of any of the examples disclosed herein; d) the VLP of any of the examples disclosed herein; or e) a combination of the foregoing wherein the *C9orf72* gene of the cells targeted by the first gNA is modified by the CasX protein. In some examples, the subject is selected from the group consisting of mouse, rat, pig, non-human primate, and human. In some examples, the *C9orf72*-related disorder is ALS or FTD. In some cases, the methods of treating a subject with a *C9orf72*-related disease result in improvement in at least one clinically-relevant parameter. In other cases, the methods of treating a subject with a *C9orf72*-related disease result in improvement in at least two clinically-relevant parameters.

[0017a] In one aspect, the present disclosure refers to a composition for use in treating a *chromosome 9 open reading frame 72 (C9orf72)*-related disorder in a subject in need thereof, wherein the use comprises: modifying a *C9orf72* gene in cells of the subject by contacting said cells with a therapeutically effective dose of: a. the system disclosed herein; b. the nucleic acid disclosed herein; c. the vector disclosed herein; or d. combinations thereof wherein the *C9orf72* gene of the cells targeted by the gNA is modified by the CasX variant protein.

[0017b] In one aspect, the present disclosure refers to use of a composition in the manufacture of a medicament for treating a *chromosome 9 open reading frame 72 (C9orf72)*-related disorder in a subject in need thereof, wherein the use comprises: modifying a *C9orf72* gene in cells of the subject by contacting said cells with a therapeutically effective dose of: a. the system disclosed herein; b. the nucleic acid disclosed herein; c. the vector disclosed herein; or d. combinations thereof wherein the *C9orf72* gene of the cells targeted by the gNA is modified by the CasX variant protein.

INCORPORATION BY REFERENCE

[0018] 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. The contents of WO 2020/247882, filed on June 5, 2020, and U.S. provisional applications