



(51) International Patent Classification:  
C12N 15/113 (2010.01)

(21) International Application Number:  
PCT/US2023/062301

(22) International Filing Date:  
09 February 2023 (09.02.2023)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
63/308,221 09 February 2022 (09.02.2022) US

(71) Applicant: **LOCANABIO, INC.** [US/US]; 3545 John Hopkins Court, Suite 200, San Diego, California 92121 (US).

(72) Inventors: **BATRA, Ranjan**; c/o Locanabio, Inc., 3545 John Hopkins Court, Suite 200, San Diego, California 92121 (US). **LARDELLI MARKMILLER, Rea**; c/o Locanabio, Inc., 3545 John Hopkins Court, Suite 200, San Diego, California 92121 (US). **ROTH, Daniela**; c/o Locanabio, Inc., 3545 John Hopkins Court, Suite 200, San Diego, California 92121 (US). **TA, Angeline**; c/o Locanabio, Inc., 3545 John Hopkins Court, Suite 200, San Diego, California 92121 (US).

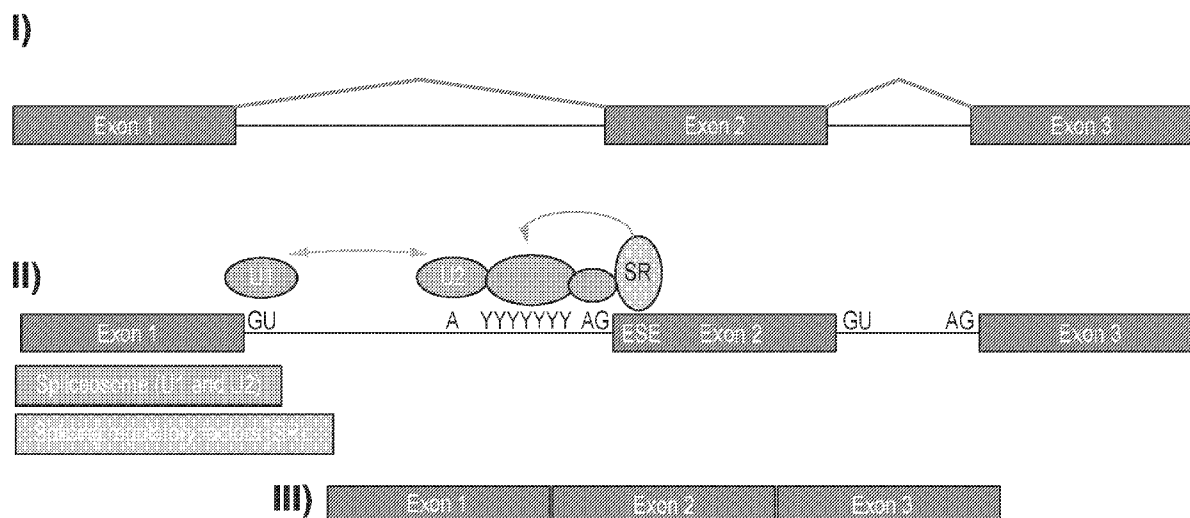
(74) Agent: **HENDERSON, Andrew** et al.; 1299 Pennsylvania Avenue, NW, Suite 700, Washington, District of Columbia 20004-2400 (US).

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

(54) Title: COMPOSITIONS AND METHODS FOR MODULATING PRE-MRNA SPLICING

FIG. 1A



(57) Abstract: Compositions capable of targeting intronic and exonic pre-mRNA sequences are disclosed that modulate mRNA splicing by excluding or including the bound pre-mRNA sequence into the mature mRNA.

**Declarations under Rule 4.17:**

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

**Published:**

- *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*
- *with sequence listing part of description (Rule 5.2(a))*

## **COMPOSITIONS AND METHODS FOR MODULATING PRE-MRNA SPLICING**

### **CROSS-REFERENCE TO RELATED APPLICATIONS**

[01] This application claims priority to, and the benefit of, U.S. Provisional Application No. 63/308,221 filed February 9, 2022. The contents of which is hereby incorporated by reference in its entirety.

### **FIELD OF THE DISCLOSURE**

[02] The disclosure is directed to molecular biology, gene therapy, and compositions and methods for modifying expression and activity of RNA molecules.

### **INCORPORATION BY REFERENCE OF SEQUENCE LISTING**

[03] The contents of the electronic sequence listing (LOCN\_013\_001WO\_SeqList\_ST26.xml; Size: 11,599,799 bytes; and Date of Creation: February 8, 2023) are herein incorporated by reference in its entirety..

### **BACKGROUND**

[04] Many human diseases arise from mutations in protein-coding genes resulting in depleted protein levels or expressed protein that lack partial or full activity. Splicing mutations, wherein a mutation that inserts, deletes, or alters the number of nucleotides at a site of RNA splicing during RNA processing, account for a third of all monogenic diseases. RNA splicing, wherein precursor messenger RNA (pre-mRNA) is processed into mature messenger RNA (mRNA) via the removal of intron sequences and the joining of exon sequences, is essential for the production of functional proteins. As such, there is an unmet need in the art to provide methods capable of altering and modulating RNA splicing in order to effectively treat monogenic disorders through the correction of splicing mutations leading to the expression of healthy and functional proteins.

[05] Examples of monogenic disorders where splicing mutations, mutations within the spliceosome, and/or frameshift mutations lead to diseases include, but are not limited to, Usher syndrome type IIA, autosomal recessive retinitis pigmentosa (RP39), Duchenne muscular dystrophies (DMD), Becker Muscular Dystrophy (BMD), amyotrophic lateral sclerosis (ALS), Alzheimer's disease, Leber congenital amaurosis (LCA), Early infantile epileptic encephalopathy-6 (EIEE6), Spinal muscular atrophy (SMA), and Huntington's

disease (HD), Dravet Syndrome, Frontal Temporal Dementia (MAPT), Parkinson Disease (PINK1). In these cases, mutations in splicing regions, including exonic and intronic sequences, lead to improper mRNA processing leading to disease. Splicing modulation can resolve many of these disorders. In addition to splicing mutations, inducing exon inclusion/exclusion can be used to resolve mutations that introduce frameshifts and missense mutations.

[06] To modulate RNA splicing, sequence specific compositions were engineered to target exonic and intronic sequences having disease-causing mutations. These compositions can modulate RNA splicing, in some cases causing mutation-bearing exons to be skipped, leading to alternately processed mRNA that, when expressed, at least partially restores the function and activity of the disease-implicated protein.

[07] Accordingly, the disclosure provides gene therapy compositions and methods comprising RNA-targeting compositions capable of modulating RNA splicing. In addition, methods of using RNA-targeting compositions to treat subjects with diseases and disorders linked to improper splicing wherein splicing modulation can be used to restore function are provided herein.

## SUMMARY

[08] The disclosure provides a pre-mRNA sequence binding composition comprising a nucleic acid sequence encoding (a) an RNA-binding polypeptide or portion thereof, and (b) one or more cognate nucleic acid guide RNA, wherein each of the one or more cognate nucleic acid guide RNA are capable of binding a target pre-mRNA sequence, wherein the target pre-mRNA sequence comprises an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), o a polypyrimidine tract (PPY), exon splicing silencer (ESS) motif, or an intron splicing silencer (ISS) motif.

[09] The disclosure provides a pre-mRNA sequence binding composition comprising a nucleic acid sequence encoding a PUF or PUMBY polypeptide capable of binding a target pre-mRNA sequence; wherein the target pre-mRNA sequence comprises an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), o a polypyrimidine tract (PPY), exon splicing silencer (ESS) motif, or an intron splicing silencer (ISS) motif.

[010] In some embodiments, the 5' splice site is a splice donor site. In some embodiments, the 3' splice site is a splice acceptor site.

**[011]** In some embodiments, the composition comprises at least two cognate nucleic acid guides.

**[012]** In some embodiments, wherein the polypeptide of is a fusion protein. In some embodiments, wherein the fusion protein comprises a peptide capable of modulating splicing. In some embodiments, wherein the peptide capable of modulating splicing binds RNA. In some embodiments, wherein the peptide capable of modulating splicing is RBFOX1, RMP38, hnRNPA1, or AF2.

**[013]** In some embodiments, the RNA-binding polypeptide is a catalytically inactivated Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated (Cas) polypeptide. In some embodiments, the Cas polypeptide is a catalytically inactivated or dead Cas13d (dCas13d) polypeptide. In some embodiments, the Cas13d polypeptide comprises an amino acid sequence set forth in SEQ ID NO: 396, SEQ ID NO: 397, SEQ ID NO: 398, SEQ ID NO: 399, SEQ ID NO: 400, SEQ ID NO: 401, or SEQ ID NO: 12307. In some embodiments, the Cas13d polypeptide comprises an amino acid sequence set forth in SEQ ID NO: 396 or SEQ ID NO: 397.

**[014]** In some embodiments, the PUF or PUMBY comprises an amino acid sequence set forth in SEQ ID NO: 380, SEQ ID NO: 381, SEQ ID NO: 382, SEQ ID NO: 383, SEQ ID NO: 384, SEQ ID NO: 385, SEQ ID NO: 386, SEQ ID NO: 387, SEQ ID NO: 388, or SEQ ID NO: 389.

**[015]** In some embodiments, the target pre-mRNA sequence comprises a sequence encoding for USH2A. In some embodiments, the sequence encoding for USH2A comprises a mutation. In some embodiments, the mutation in USH2A occurs in exon 13. In some embodiments, the pre-mRNA sequence comprises SEQ ID NO: 467. In some embodiments, the cognate nucleic acid guide RNA comprises any one of SEQ ID NOs 403-412.

**[016]** In some embodiments, the target pre-mRNA sequence comprises a sequence encoding for DMD. In some embodiments, the sequence encoding for DMD comprises a mutation. In some embodiments, the mutation in DMD occurs in at least one of exon 42, exon 43, exon 44, exon 45, exon 46, exon 47, exon 48, exon 49, exon 50, exon 51, exon 52, exon 53, exon 54, and exon 55. In some embodiments, the pre-mRNA sequences comprises SEQ ID NO: 471. In some embodiments, the cognate nucleic acid guide RNA comprises any one of SEQ ID NOs 422-437, 12272-12276, 12285, 12288-12306, or 12311-12312.

**[017]** In some embodiments, the target pre-mRNA sequence comprises a sequence encoding for MAPT. In some embodiments, the sequence encoding for MAPT comprises a mutation. In some embodiments, the mutation in MAPT occurs in exon 10. In some

embodiments, the pre-mRNA sequences comprises SEQ ID NO: 479. In some embodiments, the cognate nucleic acid guide RNA comprises any one of SEQ ID NOs 438-440.

**[018]** In some embodiments, a PUMBY protein binds a MAPT exon 10 target sequence set forth in any one of SEQ ID NO: 390, SEQ ID NO: 391, SEQ ID NO: 392, SEQ ID NO: 393, of SEQ ID NO: 394.

**[019]** In some embodiments, the nucleic acid sequence comprises at least one promoter. In some embodiments, the at least one promoter is a constitutive promoter or a tissue-specific promoter. In some embodiments, the at least one promoter is selected from the group consisting of a tCAG promoter, an EFS promoter, and an EFS-UBB promoter.

**[020]** In some embodiments, the nucleic acid sequence comprises two promoters.

**[021]** In some embodiments, a first promoter controls expression of the one or more cognate nucleic acid guide RNA and a second promoter controls expression of the RNA-binding polypeptide or portion thereof.

**[022]** The disclosure provides a vector comprising the composition of any embodiment of the disclosure. In some embodiments, the vector is selected from the group consisting of: adeno-associated virus, retrovirus, lentivirus, adenovirus, nanoparticle, micelle, liposome, lipoplex, polymersome, polyplex, and dendrimer.

**[023]** The disclosure provides a cell comprising the vector of any embodiment of the disclosure.

**[024]** The disclosure provides a method of modulating pre-mRNA splicing comprising administering a composition comprising a nucleic acid sequence encoding (a) an RNA-binding polypeptide or portion thereof, and (b) one or more cognate nucleic acid guides, wherein each of the one or more cognate nucleic acid guides are capable of binding a target pre-mRNA sequence.

**[025]** The disclosure provides a method of modulating pre-mRNA splicing comprising administering a composition comprising a nucleic acid sequence encoding a non-guided RNA-binding polypeptide or portion thereof comprising a PUF or PUMBY protein capable of binding a target pre-mRNA sequence.

**[026]** In some embodiments, the pre-mRNA sequence comprises an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), or a polypyrimidine tract (PPY). In some embodiments, the 5' splice site is a splice donor site. In some embodiments, the 3' splice site is a splice acceptor site. In some embodiments, the pre-mRNA sequence comprises an exon splicing silencer (ESS) motif or an intron splicing silencer (ISS) motif.

[027] In some embodiments, binding the pre-mRNA sequence prevents inclusion of the bound sequence into an mRNA sequence. In some embodiments, binding the pre-mRNA sequence promotes inclusion of the bound sequence into an mRNA sequence. In some embodiments, binding the pre-mRNA sequence produces a truncated mRNA sequence. In some embodiments, the mRNA sequence is expressed into a protein.

[028] In some embodiments, the protein expressed from the mRNA is expressed at an increased level compared to a wild-type mRNA. In some embodiments, the protein expressed from the mRNA is expressed at a decreased level compared to a wild-type mRNA. In some embodiments, the protein expressed from the protein has increased activity compared to a wild-type mRNA. In some embodiments, the protein expressed from the protein has decreased activity compared to a wild-type mRNA.

[029] The disclosure provides a method of treating a disease or disorder associated with a gene mutation, wherein the mutation occurs in an exon of a pre-mRNA sequence, comprising administering a composition of any embodiment of the disclosure.

[030] In some embodiments, the target pre-mRNA sequence comprises a sequence encoding for USH2A. In some embodiments, the sequence encoding for USH2A comprises a mutation. In some embodiments, the mutation in USH2A occurs in exon 13.

[031] In some embodiments, the target pre-mRNA sequence comprises a sequence encoding for DMD. In some embodiments, the sequence encoding for DMD comprises a mutation. In some embodiments, the mutation in DMD occurs in at least one of exon 42, exon 43, exon 44, exon 45, exon 46, exon 47, exon 48, exon 49, exon 50, exon 51, exon 52, exon 53, exon 54, and exon 55.

[032] In some embodiments, the disease or disorder is Duchenne muscular dystrophy, amyotrophic lateral sclerosis, or Usher syndrome type II.

[033] Any of the above aspects, or any other aspect described herein, can be combined with any other aspect.

[034] Unless otherwise defined, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure belongs. In the Specification, the singular forms also include the plural unless the context clearly dictates otherwise; as examples, the terms “a,” “an,” and “the” are understood to be singular or plural and the term “or” is understood to be inclusive. By way of example, “an element” means one or more element.

[035] Although methods and materials similar or equivalent to those described herein can be used in the practice or testing of the present disclosure, suitable methods and materials are

described below. All publications, patent applications, patents, and other references mentioned herein are incorporated by reference in their entirety. The references cited herein are not admitted to be prior art to the claimed invention. In the case of conflict, the present Specification, including definitions, will control. In addition, the materials, methods, and examples are illustrative only and are not intended to be limiting. Other features and advantages of the disclosure will be apparent from the following detailed description and claims.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

[036] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the Office upon request and payment of the necessary fee.

[037] FIG. 1A is a schematic overview of mRNA splicing depicting: I) a precursor mRNA (pre-mRNA) comprising three exonic sequences and two corresponding intron sequences linking the exons; II) in order to be processed the pre-mRNA is spliced by a protein complex comprising the spliceosome, including U1 and U2, and splicing regulator factors (SR); III) splicing removed the two intronic sequences leading to a mature mRNA that can then be translated in the cell.

[038] FIG. 1B (I-II) is a schematic depiction of the use of CRISPR-Cas nucleoprotein complexes to target specific exonic and intronic sequences to block components of the splicing machinery leading to alternative splicing.

[039] FIG. 2 is a schematic depicting a mutated exon 13 in USH2A leading to improper splicing and the expression of no protein or non-functional protein leading to Usher's syndrome type IIA. Specific targeting of exon 13 of USH2A leads to the skipping of exon 13 during mRNA processing producing a shortened, but functional Usherin protein, resulting in the treatment and correction of Usher's syndrome type IIA.

[040] FIG. 3A is a schematic showing exons 12, 13, and 14 of a human USH2A minigene pre-mRNA. Also depicted are CRISPR-Cas guide RNAs, g1-g8, targeting distinct regions of the pre-mRNA including exon 13 and intronic sequences. An exemplary guide, g7, targets the exon splicing enhancer (ESE) motif sequence of exon 13.

[041] FIG. 3B is a tape station image depicting the effect of human USH2A pre-mRNA targeting dCasRX complexes on USH2A minigene splicing, wherein the various complexes each bear guides g1-g8. The gel depicts an 860 basepair (bp) band corresponding to a

transcript comprising exons 12, 13, and 14. A second 208 bp band is depicted demonstrating the skipping of exon 13. The NT (non-targeting) control lane depicts very little exon 13 skipping, whereas each of guides g1-g8 lead to amplified exon 13 skipping.

[042] FIG. 3C is a bar graph depicting the fraction of human USH2A exon 13 excluded by dCasRX complexes comprising guides g1-g8. NT is the non-targeting control. Guide g7 results in almost complete exclusion or skipping of exon 13.

[043] FIG. 3D shows dose-dependent exon 13 skipping upon delivery of USH2A pre-mRNA targeting dCas13d complexes tested with guides g7, g9 and g10 targeting human USH2A exon 13, showing dose dependent Exon 13 skipping.

[044] FIG. 3E is a series of tape station images depicting exclusion of exon 13 of human USH2A using vectors expressing guide RNA, g7, g9, g10, or a non-targeting guide RNA. Unitary plasmids were transfected at varying doses and RNA was harvested at 48 hr.

[045] FIG. 4A is a schematic showing exons 12, 13, and 14 of a human USH2A minigene pre-mRNA. Also depicted is CRISPR-Cas guide RNA, g7, targeting exon 13.

[046] FIG. 4B is a tape station depicting the effect of human USH2A pre-mRNA targeting dCasRX complexes on USH2A minigene splicing, wherein the various complexes each include guide g7. The gel depicts an 860 base pair (bp) band corresponding to a transcript comprising exons 12, 13, and 14. A second 208 bp band is depicted demonstrating the skipping of exon 13.

[047] FIG. 5A is a bar graph depicting cas13d activity of Cas13d variants bearing point mutations to inactivate or diminish their catalytic activity. Activity is assessed by measuring relative levels of fluorescence. Cas13d variants were tested with a guide targeting a GFP reporter and relative levels of GFP were evaluated compared to a control non-targeting guide RNA.

[048] FIG. 5B is a bar graph depicting cas13d activity of active Cas13d variants to serve as a control to the inactivated variants in FIG. 5A. Activity is assessed by measuring relative levels of fluorescence. Cas13d variants with a guide targeting a GFP reporter and relative levels of GFP were evaluated compared to a control non-targeting guide RNA

[049] FIG. 6A is a schematic that depicts the targeting sequences of guides g9 and g10, which partially overlap with guide g7 and also target the ESE of exon 13 of USH2A.

[050] FIG. 6B is a tape station image depicting that AAV vector constructs expressing dSeq212 and guides g7, g9 or g10, targeting the ESE of exon 13 co-transfected with the human USH2A minigene are functional in excluding exon 13 of USH2A minigene. After 48 hours, RNA was extracted, cDNA was synthesized and PCR was performed as in FIG 3.

[051] FIG. 7A is a schematic of a mouse USH2A minigene pre-mRNA and mouse guide RNAs, mg1 and mg4-mg9, that target different sites of exon 12 of the mouse USH2A as well as intronic sequences.

[052] FIG. 7B is a tape station image depicting the effect of mouse USH2A pre-mRNA targeting dCasRX complexes on USH2A minigene splicing, wherein the various complexes each bear guides mg4-mg9. The gel depicts a 770 basepair (bp) band corresponding to a transcript comprising exons 11, 12, and 13. A second approximately 350 bp band is depicted corresponding to partial exon 12 exclusion. A third 128 bp band is depicted demonstrating the skipping of exon 12.

[053] FIG. 8A is a tape station image depicting exon 12 skipping of mouse USH2A using AAV vector comprising Cas13d-expression constructs that co-express both mg1 and mg7 in tandem array.

[054] FIG. 8B is a series of microscope images depicting splice junctions of mouse USH2A minigene in cosm6 cells transfected with 250 ng of guides mg1 and mg7, 500 ng dCasRx and 50 ng mouse minigene in 12 well format. Cells were trypsinized and seeded in Matrigel-coated chamber slides 24 hr before fixation with PFA. Cells not treated with minigene show no signal from either set of probes, cells treated with a non-targeting (NT) guide show signal only from exon 11/exon12 splice junction, cells treated with both mg1 and mg7 show reduced signal for the exon 11/exon 12 splice junction and presence of the signal from exon 11/exon 13 splice junction (indicative of exon 12 skipping).

[055] FIG. 9A is a schematic depicting an exemplary AAV vector construct encoding mg1 and mg7 guide RNAs under the control of a human U6 (hU6) promoter and a dCas13d-NLS (dCas13d tethered to a nuclear localization signal) under the control of an EFS promoter for the targeting of exon 12 of mouse USH2A.

[056] FIG. 9B is a schematic depicting an exemplary AAV vector construct encoding a g7 guide RNA under the control of a human U6 (hU6) promoter and a dCas13d-NLS (dCas13d tethered to a nuclear localization signal) under the control of a cone/rod specific promoter for the targeting of exon 13 of human USH2A.

[057] FIG. 10 is a series of tape station images depicting exclusion of exon 12 of mouse USH2A using vectors expressing dual guide RNA, mg1 and mg7. Unitary plasmids were constructed expressing dSeq212 under EFS or EFS-UBB promoters and mg1 and mg7 either under separate hU6 promoters or as a tandem array. Unitary plasmids were transfected at varying doses with the mouse USH2A minigene and RNA was harvested at 48 hr. Splicing assays were performed as described above.

[058] FIG. 11A is a schematic depicting the targeting of exon 51 of endogenous DMD with guide RNA.

[059] FIG. 11B a series of tape station images depicting the impact of guides g2-g7 with either dCasRx or dSeq189 on exon 51 skipping. The gel image depicts a 322 bp band corresponding to a transcript comprising exons 50, 51 and 52. A second 89 bp band is depicted corresponding to a transcript comprising exons 50 and 52, having skipped exon 51.

[060] FIG. 12 is a schematic depicting a multi-exon skipping approach to restore dystrophin levels by skipping exons 45-50 or exons 46-50.

[061] FIG. 13A is a schematic depicting guides that target exons 45 and 51 of DMD.

[062] FIG. 13B is a tape station image depicting DMD exon skipping following treatment with dCasRX and guide RNAs targeting exons 45 and 51.

[063] FIG. 14 is a tape station image image depicting the impact of guides g1-g3, targeting exon 10 of a human MAPT minigene, complexed with dCasRx. The image depicts a 197 bp band corresponding to a transcript comprising exon 10. A second 104 bp band is also depicted corresponding to the absence of exon 10, indicative of exon 10 skipping. JO and JP are MAPT minigenes for exon 10 skipping.

[064] FIG. 15 is tape station image depicting the impact of targeting exon 10 of a human MAPT minigene with PUMBY proteins. A positive control is dCasRx complexed with guide g3 from FIG. 14. The gel image depicts a 197 bp band corresponding to a transcript comprising exon 10. A second 104 bp band is also depicted corresponding to the absence of exon 10, indicative of exon 10 skipping. JO is a MAPT minigene for exon 10 skipping.

[065] FIG. 16A is a schematic depicting a strategy to knockdown a transcript such as SOD1 by introduction of a PTC (premature termination codon) through exon skipping. The image shows SOD1 cDNA comprising exons 1-5. Orange reverse arrows demonstrate where guides target exon sequences of SOD1. Forward arrows depict predicted enhancer sequences within the exon sequences.

[066] FIG. 16B is a schematic depicting exons 1-5 of endogenous SOD1 and guide RNAs targeting exon 2 and exon 3.

[067] FIG. 16C is a tape station image depicting the impact of guides g1-g4, targeting exons 2 and 3 of endogenous SOD1, complexed with dCasRx. The gel depicts a 224 bp band corresponding to a transcript comprising exons 1-4. A second 154 bp band is also depicted corresponding to the absence of exon 3, indicative of exon 3 skipping. A third 127 bp band is also depicted corresponding to the absence of exon 2, indicative of exon 2 skipping. A fourth 57 bp band is also depicted corresponding to the absence of exons 2 and 3, indicative of exon

2 and 3 skipping. Skipping of exons 2 and/or 3 is predicted to lead to NMD (nonsense mediated decay).

[068] FIG. 17 is a schematic depicting an LRRK3 cDNA comprising exons 1-4.

[069] FIG. 18A is an exemplary vector schematic expressing green fluorescent protein (GFP).

[070] FIG. 18B is a series of immunofluorescent microscope images of mice retinal cryosections following subretinal injection of AAV5, 8, 9 or Rh10 vectors encoding GFP.

[071] FIG. 18C is a vector schematic depicting a AAV8-U6-mg1/mg7-dCAs13d vector. U6 = human U6 promoter; mg1 = mouse USH2A exon 13 guide 1; mg7 = mouse USH2A exon 13 guide 7.

[072] FIG. 18D is a series of immunofluorescent RNAscope microscope images and BaseScope images of mice retinal cryosections injected with control vector or AAV8-U6-mg1/mg7-dCAs13d vector. RNAscope probes are specific for USH2A. BaseScope probes were specific for spliced (Exon11/13) and intact (Exon11/12) mRNA

[073] FIG. 18E is a bar graph depicting percent exon 12 exclusion following treatment with control vector or AAV8-U6-mg1/mg7-dCAs13d vector.

[074] FIG. 19A is vector schematic depicting a AAV8-U6-mg1/mg7-dCAs13d vector. U6 = U6 promoter; mg1 = mouse USH2A exon 12 guide 1; mg7 = mouse USH2A exon 12 guide 7.

[075] FIG. 19B is a series of immunofluorescent microscope images. Cone photoreceptors were immunolabeled with the anti-cone opsin Antibody (Ab). Rod photoreceptors were immunolabeled with the anti-rhodopsin Ab. T Cells were immunolabeled using the anti-CD3 Ab.

[076] FIG. 20 is a schematic depicting exon skipping of exon 51 in DMD.

[077] FIG. 21 is a schematic depicting that therapies designed to skip multiple exons may benefit around 65% of DMD patients.

[078] FIG. 22A is a schematic showing exon 51 of a human DMD pre-mRNA and CRISPR-Cas guide RNAs targeting distinct regions of the pre-mRNA including ESE sequences.

[079] FIG. 22B is a series of tape station images depicting the impact of guides from FIG. 22A with dSeq212/4aa mutation on exon 51. The gel image depicts a band corresponding to a transcript comprising exon 51. A second band is depicted corresponding to a transcript having skipped exon 51.

[080] FIG. 22C is a graph depicting the percentage of exon 51 excluded for different exon 51 targeting guides.

[081] FIG. 23A schematic showing exon 51 of a human DMD pre-mRNA and CRISPR-Cas guide RNAs targeting distinct regions of the pre-mRNA including ESE sequences.

[082] FIG. 23B is a series of tape station images depicting the impact of multiple guides from FIG. 23A with dSeq212/4aa mutation on exon 51. In some cases, multi-targeting was performed with two guides targeting distinct sites of exon 51. The gel image depicts a band corresponding to a transcript comprising exon 51. A second band is depicted corresponding to a transcript having skipped exon 51.

[083] FIG. 23C is a graph depicting percentage of exon 51 excluded for different exon 51 targeting guides alone or in combination with another exon 51 guide.

[084] FIG. 24 is a tape station image and a diagram depicting additional DMD exons for targeting to induce further DMD exon skipping.

[085] FIG. 25A is a tape station image depicting the impact of A03278 AAV\_hU6-36DR-g38-36DR-g42(tandem)\_CMV-dSeq212-4aa-NLS-WPRE3 (dCas13d dSeq212 complexed with guides g38 and g42), targeting exon splice enhancers in DMD exon 51 in human myotubes with a deletion of exon 52. The deletion of exon 52 results in no dystrophin protein expression. A 496 bp band represents the exon 51 included DMD transcript. A 263 bp band corresponds to the absence of exon 51 and 52, indicative of exon 51 and 52 skipping.

[086] FIG. 25B is a graph depicting percentage of exon 51 skipping upon treatment with A03278.

[087] FIG. 25C is a series of immunofluorescence images depicting dystrophin expression. Restoration of dystrophin expression is observed upon A03278 treatment compared to an untreated control which expresses no dystrophin protein due to deletion of exon 52.

[088] FIG. 26A is a schematic showing exon 7 and intron 7-8 of a human SMN2 minigene pre-mRNA and guides g1-g5 targeting the intronic region.

[089] FIG. 26B is an SMN2 minigene comprising exon 6, exon 7, and exon 8. Guides from FIG. 26A are depicted targeting the intronic ISS sequence.

[090] FIG. 26C is a tape station image depicting the impact of guides g1-g5 targeting the intronic ISS between exons 7 and 8 of SMN2, complexed with dCas13d, on mRNA splicing. A 218bp band is depicted corresponding to exon 6 and exon 8, with exon 7 being skipped. A second 272 bp band is also depicted corresponding to the exons 6, 7, and 8 indicative of the inclusion of exon 7.

[091] FIG. 26D is a graph depicting percentage of exon 7 inclusion upon treatment with dCas13d and multiple guides described in FIG.26A.

#### DETAILED DESCRIPTION

[092] The disclosure provides gene therapy compositions comprising polypeptides, polynucleotides, or nucleoprotein complexes capable of binding target pre-mRNA sequences and modulating RNA splicing and mRNA processing.

[093] Disclosed herein are vectors comprising the pre-mRNA binding compositions of the disclosure.

[094] A “pre-mRNA sequence”, or precursor mRNA, is a single-stranded RNA sequence comprising exonic and intronic sequences synthesized by transcription of DNA that has not yet undergone processing to become messenger RNA (mRNA). In some aspects, the pre-mRNA sequence comprises a mutation. In some aspects, the mutation is a base-pair change, a deletion, an insertion, or a frame-shift mutation. In some aspects, the mutation in the pre-mRNA sequence produces an mRNA sequence that encodes for a protein that is one or more of: not expressed, expressed poorly, truncated, misfolded, catalytically inactive, or catalytically hyperactive. In some aspects, the mutation occurs in an exonic region.

[095] In some aspects, the pre-mRNA sequence comprises at least one of: an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), and a polypyrimidine tract (PPY). In some aspects, the 5' splice site is a splice donor site. In some aspects, the 3' splice site is a splice acceptor site. In some aspects, the pre-mRNA sequence comprises an exon splicing silencer (ESS) motif or an intron splicing silencer (ISS) motif.

[096] Messenger RNA processing occurs in the nucleus and comprises modifying pre-mRNA to include 5' capping, 3' polyadenylation, and RNA splicing. RNA splicing involves removing intronic sequences, which do not code for proteins, from the transcript and re-joining the exonic sequences to form a mature mRNA sequence that codes for a given protein sequence that is expressed or translated in the cell.

[097] RNA splicing is performed by the spliceosome, a large RNA-protein complex comprised of small nuclear ribonucleoproteins (snRNP) which assemble at intronic 5' and 3' splice sites on the pre-mRNA to remove intronic sequences and ligate the exonic sequences to form the mature mRNA sequence (FIG. 1A). Components of the spliceosome include, but are not limited to, U1, U2, U4, U5, and U6 snRNPs. Further components include splicing

regulatory factors such as splicing factor 1 (SF1), splicing factor U2AF 35 kd (U2AF35), and splicing factor U2AF 65 kd (U2AF65).

[098] Compositions of the disclosure, and methods of using the compositions, comprise sequence specific RNA binding proteins, nucleic acids, and nucleoprotein complexes, that bind intronic and exonic sequences of pre-mRNA sequences. The binding of the pre-mRNA by these sequences blocks, inhibits, or prevents components of the spliceosome from assembling at the blocked site thereby preventing inclusion of the bound exon into the mature mRNA sequence. The resultant modified mRNA sequence lacks the skipped exon or exons. As such, the transcribed mRNA can produce a polypeptide lacking the amino acids that are encoded by the skipped exon or exons.

[099] In some aspects, compositions of the disclosure, and methods of using the compositions, comprise sequence specific RNA binding proteins, nucleic acids, and nucleoprotein complexes, that bind intronic and exonic sequences of pre-mRNA sequences. The binding of the pre-mRNA by these sequences promotes inclusion of exons into mature mRNA sequences that are otherwise skipped during processing. In some embodiments, exon inclusion is promoted by specifically binding an exon splicing silencer (ESS) motif or an intron splicing silencer (ISS) motif.

[0100] In some aspects, RNA binding proteins, nucleic acids, and nucleoprotein complexes of the disclosure bind a pre-mRNA sequence comprising at least one of an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), a polypyrimidine tract (PPY), an exon splicing silencer (ESS) motif, or an intron splicing silencer (ISS) motif.

#### **RNA-Guided RNA-Binding Systems**

[0101] In some embodiments, the disclosure provides RNA-guided RNA-binding systems. In some aspects, the RNA-guided RNA-binding system is an RNase Cas-based RNA-guided RNA-binding polypeptide. In some embodiments, a nucleic acid sequence encodes an RNA-guided RNA-binding polypeptide which is an RNase Cas protein (or a deactivated RNase Cas protein). In one embodiment, the nucleic acid sequence further comprises a gRNA sequence comprising a spacer sequence which binds to a pre-mRNA sequence of the disclosure and a direct repeat (DR) sequence which binds to the RNase Cas protein.

[0102] In one embodiment, a Cas13d system is catalytically active, in which case, the Cas13d nucleoprotein complex cleaves and destroys a pre-mRNA sequence. In another embodiment, a Cas13d system is catalytically inactive, in which case, the Cas13d nucleoprotein complex binds the pre-mRNA sequence. In another embodiment, the

catalytically inactive Cas13d nucleoprotein complex binds the pre-mRNA sequence and modifies pre-mRNA splicing and processing. In some aspects, this splicing modification results in the inclusion or exclusion of one or more exons in the mature mRNA sequence. In yet another embodiment, a Cas13d comprises a catalytically inactive Cas13d fused to an endonuclease which is capable of cleaving the target pre-mRNA or mRNA sequence. In such an embodiment, the endonuclease is an active RNase. Exemplary endonucleases with RNase activity can be found herein, and these include, for example, a domain from a ZC3H12A zinc-finger (also referred herein as E17) or a PIN endonuclease.

**[0103]** In one embodiment, the RNase Cas protein is a Cas13 protein. In another embodiment, the Cas13 protein is a Cas13d protein. In another embodiment, the Cas13d protein is a deactivated RNase Cas13d protein (dCas13d). In another embodiment, the dCas13d protein is a fusion protein comprising 1) dCas13d and 2) a polypeptide encoding a protein or fragment thereof having nuclease activity. In another embodiment, the dCas13d protein is a fusion protein comprising 1) dCas13d and 2) a nuclease domain of ZC3H12A, a zinc-finger endonuclease, (referred to as E17 herein). In some embodiments, the Cas configuration comprises a signal sequence(s) such as NLS(s) and/or NES(s). In some embodiments, the dCas13d is linked to E17 via a linker sequence. In one embodiment, the linker sequence is VDTANGS. In some embodiments, the nucleic acid sequence encoding the Cas13d or dCas13d fusion proteins are operably linked to at least one promoter sequence. In some embodiments, the promoter sequence comprises an enhancer and/or an intron. In some embodiments, the promoter sequence is a constitutive promoter such as, EIF1a or its truncated form, the EFS promoter sequence, or the full length or truncated (tCAG) promoter sequence or the EFS/UBB promoter sequence. In other cases the promoter is a tissue specific promoter such as the neuron specific synapsin promoter sequence. In some embodiments, the nucleic acid sequence comprises a first promoter sequence that controls expression of a Cas13d protein or Cas13d fusion protein and a second promoter sequence that controls expression of the at least one guide RNA sequence.

**[0104]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA-guided RNA binding protein comprises a sequence isolated or derived from a protein with no DNA nuclease activity.

**[0105]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA-guided RNA binding protein disclosed herein comprises a sequence isolated or derived from a CRISPR Cas protein. In some embodiments, the CRISPR Cas protein is not a

Type II CRISPR Cas protein. In some embodiments, the CRISPR Cas protein is not a Cas9 protein.

**[0106]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA-guided RNA binding protein comprises a Type VI CRISPR Cas protein or portion thereof. In some embodiments, the Type VI CRISPR Cas protein comprises a Cas13 protein or portion thereof. Exemplary Cas13 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, bacteria or archaea. Exemplary Cas13 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, *Leptotrichia wadei*, *Listeria seeligeri* serovar 1/2b (strain ATCC 35967 / DSM 20751 / CIP 100100 / SLCC 3954), *Lachnospiraceae* bacterium, *Clostridium aminophilum* DSM 10710, *Carnobacterium gallinarum* DSM 4847, *Paludibacter propionigenes* WB4, *Listeria weihenstephanensis* FSL R9-0317, *Listeria weihenstephanensis* FSL R9-0317, bacterium FSL M6-0635 (*Listeria newyorkensis*), *Leptotrichia wadei* F0279, *Rhodobacter capsulatus* SB 1003, *Rhodobacter capsulatus* R121, *Rhodobacter capsulatus* DE442 and *Corynebacterium ulcerans*. Exemplary Cas13 proteins of the disclosure may be DNA nuclease inactivated. Exemplary Cas13 proteins of the disclosure include, but are not limited to, Cas13a, Cas13b, Cas13c, Cas13d and orthologs thereof. Exemplary Cas13b proteins of the disclosure include, but are not limited to, subtypes 1 and 2 referred to herein as Csx27 and Csx28, respectively.

**[0107]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA binding protein comprises a sequence isolated or derived from a Cas13d protein. Cas13d is an effector of the type VI-D CRISPR-Cas systems. In some embodiments, the Cas13d protein is an RNA-guided RNA endonuclease enzyme that can cut or bind RNA. In some embodiments, the Cas13d protein can include one or more higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domains. In some embodiments, the Cas13d protein can include either a wild-type or mutated HEPN domain. In some embodiments, the Cas13d protein includes a mutated HEPN domain that cannot cut RNA but can process guide RNA. In some embodiments, the Cas13d protein does not require a protospacer flanking sequence. Also see WO Publication No. WO2019/040664 & US2019/0062724, which is incorporated herein by reference in its entirety, for further examples and sequences of Cas13d protein, without limitation.

**[0108]** In some embodiments, Cas13d sequences of the disclosure include without limitation SEQ ID NOS: 1-296 of WO 2019/040664, so numbered herein and included herewith. Yan et al. (2018) Mol Cell. 70(2):327-339 (doi: 10.1016/j.molcel.2018.02.2018)

and Konermann et al. (2018) Cell 173(3):665-676 (doi: 10.1016/j.cell/2018.02.033) have described Cas13d proteins and both of which are incorporated by reference herein in their entireties. Also see WO Publication Nos. WO2018/183403 (CasM, which is Cas13d) and WO2019/006471 (Cas13d), which are incorporated herein by reference in their entirety.

### **Guide RNAs for RNA-Guided RNA-Binding Proteins**

[0109] The terms guide RNA (gRNA), cognate nucleic acid guide, and single guide RNA (sgRNA) are used interchangeably throughout the disclosure.

[0110] Guide RNAs (gRNAs) of the disclosure may comprise a spacer sequence and a “direct repeat” (DR) sequence. In some embodiments, a guide RNA is a single guide RNA (sgRNA) comprising a contiguous spacer sequence and DR sequence. In some embodiments, the spacer sequence and the DR sequence are not contiguous. In some embodiments, the gRNA comprises a DR sequence. DR sequences refer to the repetitive sequences in the CRISPR locus (naturally-occurring in a bacterial genome or plasmid) that are interspersed with the spacer sequences. It is well known that one would be able to infer the DR sequence of a corresponding (or cognate) Cas protein if the sequence of the associated CRISPR locus is known. In some embodiments, a guide RNA comprises a direct repeat (DR) sequence and a spacer sequence. In some embodiments, a sequence encoding a guide RNA or single guide RNA of the disclosure comprises or consists of a spacer sequence and a DR sequence, that are separated by a linker sequence. In some embodiments, the linker sequence may comprise or consist of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 or any number of nucleotides (nt) in between. In some embodiments, the linker sequence may comprise at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50 or any number of nucleotides in between. In some embodiments, the DR sequence is a Cas13d DR sequence.

[0111] In one embodiment, the gRNA that hybridizes with the one or more target RNA molecules in a Cas 13d-mediated manner includes one or more direct repeat (DR) sequences, one or more spacer sequences, such as, e.g., one or more sequences comprising an array of DR-spacer-DR-spacer. In one embodiment, a plurality of gRNAs are generated from a single array, wherein each gRNA can be different, for example target different RNAs or target multiple regions of a single RNA, or combinations thereof. In some embodiments, an isolated gRNA includes one or more direct repeat sequences, such as an unprocessed (e.g., about 36 nt) or processed DR (e.g., about 30 nt). In some embodiments, a gRNA can further include one or more spacer sequences specific for (e.g., is complementary to) the target RNA. In certain such embodiments, multiple polIII promoters can be used to drive multiple gRNAs,

spacers and/or DRs. In one embodiment, a guide array comprises a DR (about 36nt)-spacer (about 30nt)-DR (about 36nt)-spacer (about 30nt).

**[0112]** Guide RNAs (gRNAs) of the disclosure may comprise non-naturally occurring nucleotides. In some embodiments, a guide RNA of the disclosure or a sequence encoding the guide RNA comprises or consists of modified or synthetic RNA nucleotides. Exemplary modified RNA nucleotides include, but are not limited to, pseudouridine (Ψ), dihydrouridine (D), inosine (I), and 7-methylguanosine (m7G), hypoxanthine, xanthine, xanthosine, 7-methylguanine, 5, 6-Dihydrouracil, 5-methylcytosine, 5-methylcytidine, 5-hydropoxymethylcytosine, isoguanine, and isocytosine.

**[0113]** Guide RNAs (gRNAs) of the disclosure may bind modified RNA within a target sequence. Within a target sequence, guide RNAs (gRNAs) of the disclosure may bind modified or mutated (e.g., pathogenic) RNA. Exemplary epigenetically or post-transcriptionally modified RNA include, but are not limited to, 2'-O-Methylation (2'-OMe) (2'-O-methylation occurs on the oxygen of the free 2'-OH of the ribose moiety), N6-methyladenosine (m6A), and 5-methylcytosine (m5C).

**[0114]** In some embodiments of the compositions of the disclosure, a guide RNA of the disclosure comprises at least one sequence encoding a non-coding C/D box small nucleolar RNA (snoRNA) sequence. In some embodiments, the snoRNA sequence comprises at least one sequence that is complementary to the target RNA, wherein the target sequence of the RNA molecule comprises at least one 2'-OMe. In some embodiments, the snoRNA sequence comprises at least one sequence that is complementary to the target RNA, wherein the at least one sequence that is complementary to the target RNA comprises a box C motif (RUGAUGA SEQ ID NO: 375) and a box D motif (CUGA SEQ ID NO: 376).

**[0115]** Spacer sequences of the disclosure bind to the target sequence of an RNA molecule. In some embodiments, spacer sequences of the disclosure bind to pathogenic target RNA. In some embodiments, spacer sequences of the disclosure bind pre-mRNA sequences. In some embodiments, spacer sequences of the disclosure bind intronic sequences of pre-mRNA. In some embodiments, spacer sequences of the disclosure bind exonic sequences of pre-mRNA.

**[0116]** In some embodiments of the compositions of the disclosure, the sequence comprising the gRNA comprises a spacer sequence that specifically binds to the target RNA sequence. In some embodiments, the spacer sequence has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 87%, 90%, 95%, 97%, 99% or any percentage in between of complementarity to the target RNA sequence. In some embodiments, the spacer sequence has 100% complementarity to the target RNA sequence. In some embodiments, the spacer

sequence comprises or consists of 20 nucleotides. In some embodiments, the spacer sequence comprises or consists of 21 nucleotides, 22 nucleotides, 23 nucleotides, 24 nucleotides, 25 nucleotides, 26 nucleotides, 27 nucleotides, 28 nucleotides, 29 nucleotide, 30 nucleotides, 31 nucleotides or 32 nucleotides. In some embodiments, the spacer sequence comprises or consists of 26 nucleotides. In some embodiments, the spacer sequence is non-processed and comprises or consists of 30 nucleotides. In some embodiments the non-processed spacer sequence comprises or consists of 30-36 nucleotides.

**[0117]** DR sequences of the disclosure bind the Cas polypeptide of the disclosure. Upon binding of the spacer sequence of the gRNA to the target RNA sequence, the Cas protein bound to the DR sequence of the gRNA is positioned at the target RNA sequence. In some aspects, a DR sequence has at least 50%, 55%, 60%, 65%, 70%, 75%, 80%, 87%, 90%, 95%, 97%, 99%, 100% or any percentage in between identity to a DR sequence disclosed herein. In some embodiments, DR sequences of the disclosure comprise a secondary structure or a tertiary structure. Exemplary secondary structures include, but are not limited to, a helix, a stem loop, a bulge, a tetraloop and a pseudoknot. Exemplary tertiary structures include, but are not limited to, an A-form of a helix, a B-form of a helix, and a Z-form of a helix. Exemplary tertiary structures include, but are not limited to, a twisted or helicized stem loop. Exemplary tertiary structures include, but are not limited to, a twisted or helicized pseudoknot. In some embodiments, DR sequences of the disclosure comprise at least one secondary structure or at least one tertiary structure. In some embodiments, DR sequences of the disclosure comprise one or more secondary structure(s) or one or more tertiary structure(s).

**[0118]** In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof selectively binds to a tetraloop motif in an RNA molecule of the disclosure. In some embodiments, a target sequence of an RNA molecule comprises a tetraloop motif. In some embodiments, the tetraloop motif is a “GRNA” motif comprising or consisting of one or more of the sequences of GAAA, GUGA, GCAA or GAGA.

**[0119]** In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof that binds to a target sequence of an RNA molecule hybridizes to the target sequence of the RNA molecule. In some embodiments, a guide RNA or a portion thereof that binds to a first RNA binding protein or to a second RNA binding protein covalently binds to the first RNA binding protein or to the second RNA binding protein. In some embodiments, a guide RNA or a portion thereof that binds to a first RNA binding protein or to a second RNA

binding protein non-covalently binds to the first RNA binding protein or to the second RNA binding protein.

**[0120]** In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof comprises or consists of between 10 and 100 nucleotides, inclusive of the endpoints. In some embodiments, a spacer sequence of the disclosure comprises or consists of between 10 and 30 nucleotides, inclusive of the endpoints. In some embodiments, a spacer sequence of the disclosure comprises or consists of 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 nucleotides. In some embodiments, the spacer sequence of the disclosure comprises or consists of 20 nucleotides. In some embodiments, the spacer sequence of the disclosure comprises or consists of 21 nucleotides. In some embodiments, the spacer sequence of the disclosure comprises or consists of 26 nucleotides.

**[0121]** Guide molecules generally exist in various states of processing. In one example, an unprocessed guide RNA is 36nt of DR followed by 30-32 nt of spacer. The guide RNA is processed (truncated/modified) by Cas I3d itself or other RNases into the shorter "mature" form. In some embodiments, an unprocessed guide sequence is about, or at least about 30, 35, 40, 45, 50, 55, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150 or more nucleotides (nt) in length. In some embodiments, a processed guide sequence is about 40 to 150 nt (such as 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150 nt).. In some embodiments, an unprocessed spacer is about 28-40 nt long (such as 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, or 40 nt) while the mature (processed) spacer can be about 10 to 30 nt, 10 to 25 nt, 14 to 25 nt, 20 to 22 nt, or 14-30 nt (such as 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, or 35 nt). In some embodiments, an unprocessed DR is about 36 nt (such as 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40 or 41 nt), while the processed DR is about 30 nt (such as 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, or 35 nt). In some embodiments, a DR sequence is truncated by 1-10 nucleotides (such as 1, 2, 3, 4, 5, 6, 7, 8, 9, to 10 nucleotides at e.g., the 5' end in order to be expressed as mature pre-processed guide RNAs.

**[0122]** In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof does not comprise a nuclear localization sequence (NLS).

**[0123]** In some embodiments of the compositions of the disclosure, a guide RNA or a portion thereof comprises a sequence complementary to a protospacer flanking sequence

(PFS). In some embodiments, including those wherein a guide RNA or a portion thereof comprises a sequence complementary to a PFS, the first RNA binding protein may comprise a sequence isolated or derived from a Cas13 protein. In some embodiments, including those wherein a guide RNA or a portion thereof comprises a sequence complementary to a PFS, the first RNA binding protein may comprise a sequence encoding a Cas13 protein or an RNA-binding portion thereof. In some embodiments, the guide RNA or a portion thereof does not comprise a sequence complementary to a PFS.

**[0124]** In some embodiments, the spacer sequence which binds the target pre-mRNA sequence comprises or consists of about 20-40 nucleotides. In some embodiments, a gRNA comprises one or more spacer sequences.

**[0125]** Exemplary gRNA spacer sequences of the disclosure specifically bind to pre-mRNA encoding genes including USH2A (usherin), DMD (dystrophin), SOD1 (superoxide dismutase 1), LRRK2 (Leucine-rich repeat kinase 2), MAPT (microtubule-associated protein tau), CEP290 (Centrosomal protein of 290 kDa), SCN1A (sodium voltage-gated channel alpha subunit 1), SMN1 (Survival of motor neuron 1), SMN2 (survival of motor neuron 2), or HTT (huntingtin).

**[0126]** Exemplary gRNA spacer sequences of the disclosure specifically bind to a USH2A pre-mRNA. In some embodiments, the sequence encoding for USH2A comprises a mutation. In some embodiments, the USH2A pre-mRNA is human. In some embodiments, the USH2A pre-mRNA is murine. In some embodiments, the gRNA spacer sequences of the disclosure specifically bind exon 13 of human USH2A pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind the ESE motif of exon 13 of human USH2A pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind intronic sequences 5' and 3' of exon 13 of human USH2A pre-mRNA. In some aspects, USH2A gRNA spacer sequences of the disclosure are selected from those listed in Table 1. In some aspects, the gRNA spacer sequence targeting human USH2A is at least one of SEQ ID NO: 528-1413.

**Table 1: Human USH2A exon 13 gRNA spacer sequences**

| Guide | Target Site                      | Spacer Sequence                                |
|-------|----------------------------------|------------------------------------------------|
| g1    | Polypyrimidine (PPY) tract (YYY) | TTTAAAAGGTGAGGATGGGAAAATGA<br>(SEQ ID NO: 403) |

|     |                                              |                                                |
|-----|----------------------------------------------|------------------------------------------------|
| g2  | 3' splice site                               | CCCTAAAGATAAAATATATTTAAAAG<br>(SEQ ID NO: 404) |
| g3  | 3' splice site                               | CACACCTAAGCCCTAAAGATAAAATA<br>(SEQ ID NO: 405) |
| g4  | 5' splice site                               | ACATTTCTTTCTTACCTGGTTGACAC<br>(SEQ ID NO: 406) |
| g5  | 5' splice site                               | AATGTAATACATTTCTTTCTTACCTG<br>(SEQ ID NO: 407) |
| g6  | exon splicing<br>enhancer (ESE)<br>motif     | CTGAGCCATGGAGGTTACACTGGCAG<br>(SEQ ID NO: 408) |
| g7  | exon splicing<br>enhancer (ESE)<br>motif     | ACACACAGGCACTGGCCACTGATTGG<br>(SEQ ID NO: 409) |
| g8  | intronic splicing<br>enhancer (ISE)<br>motif | GATGGTAATAGAGATGTGACTGTAAA<br>(SEQ ID NO: 410) |
| g9  | exon splicing<br>enhancer (ESE)<br>motif     | AGGCACTGGCCACTGATTGGGTCACA<br>(SEQ ID NO: 411) |
| g10 | exon splicing<br>enhancer (ESE)<br>motif     | ACTGGCCACTGATTGGGTCACAAATG<br>(SEQ ID NO: 412) |

[0127] In some embodiments, the gRNA spacer sequences of the disclosure specifically bind exon 12 of murine USH2A pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind the ESE motif of exon 12 of murine USH2A pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind intronic sequences 5' and 3' of exon 12 of murine USH2A pre-mRNA. In some aspects, murine USH2A gRNA spacer sequences of the disclosure are selected from those listed in Table 2.

**Table 2: Murine USH2A exon 12 gRNA spacer sequences**

|                |                                                 |
|----------------|-------------------------------------------------|
| mg1            | AGACATAGACACTGGCCACTGATGGG (SEQ ID NO: 413)     |
| mg1 –<br>30 nt | GTAGACATAGACACTGGCCACTGATGGGGT (SEQ ID NO: 414) |
| mg4            | GCTGCAGTCACAGACCTCACAAGCAC (SEQ ID NO: 415)     |
| mg5            | CTGGAAACAAAACACAATGGAAAGGA (SEQ ID NO: 416)     |

|             |                                                 |
|-------------|-------------------------------------------------|
| mg6         | GAAATTGCAGCGATTGCATCTAAGCC (SEQ ID NO: 417)     |
| mg7         | AGCGATTGCATCTAAGCCCTGGAAAC (SEQ ID NO: 418)     |
| mg7 – 30 nt | GCAGCGATTGCATCTAAGCCCTGGAAACAA (SEQ ID NO: 419) |
| mg8         | TCTCTGCAAGAGTCACACTTAAGTCC (SEQ ID NO: 420)     |
| mg9         | AGTCCCAGGCTGGGACCCAGCCTTGC (SEQ ID NO: 421)     |

**[0128]** Exemplary gRNA spacer sequences of the disclosure specifically bind to a DMD pre-mRNA. In some embodiments, the sequence encoding for DMD comprises a mutation. In some embodiments, the DMD pre-mRNA is human. In some embodiments, the gRNA spacer sequences of the disclosure specifically bind at least one of exon 45, exon 46, exon 47, exon 48, exon 49, exon 50, and exon 51 of human DMD pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure specifically bind exon 51 of human DMD pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind the ESE motif of exon 51 of human DMD pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind intronic sequences 5' and 3' of exon 51 of human DMD pre-mRNA. In some aspects, DMD pre-mRNA gRNA spacer sequences of the disclosure are selected from those listed in Table 3. In some aspects, the gRNA spacer sequence targeting human DMD pre-mRNA is at least one of SEQ ID NO: 1414-12271.

**Table 3: Human DMD exon 51 gRNA spacer sequences**

| Guide | Spacer Sequence                                  |
|-------|--------------------------------------------------|
| g1    | ggaaaaaagaagaaaaagaaaaTTA (SEQ ID NO: 422)       |
| g2    | GGAGCTAAAATATTTTGGGTTTTTGC (SEQ ID NO: 423)      |
| g3    | GTCTGAGTAGGAGCTAAAATATTTTG (SEQ ID NO: 424)      |
| g4    | TGAGTAGGAGCTAAAATATTTTGGGT (SEQ ID NO: 425)      |
| g5    | AACATCAAGGAAGATGGCATTCTAG (SEQ ID NO: 426)       |
| g5+   | AACATCAAGGAAGATGGCATTCTAGTTTG (SEQ ID NO: 12311) |
| g6    | ACATCAAGGAAGATGGCATTCT (SEQ ID NO: 427)          |

|      |                                                   |
|------|---------------------------------------------------|
| g7   | CCTCTGTGATTTTATAACTTGAT (SEQ ID NO: 428)          |
| g8   | TCATTTTTTCTCATACCTTCTGCTTG (SEQ ID NO: 429)       |
| g9   | CTTTTATCATTTTTTCTCATACCTTC (SEQ ID NO: 430)       |
| g10  | TTTTTCTCATACCTTCTGCTTGATGA (SEQ ID NO: 431)       |
| g11  | CTGAGTAGGAGCTAAAATATTTTGGG (SEQ ID NO: 432)       |
| g12  | AGTAGGAGCTAAAATATTTTGGG (SEQ ID NO: 433)          |
| g13  | CAGAGTAACAGTCTGAGTAGGAGCTA (SEQ ID NO: 434)       |
| g14  | GTAACAGTCTGAGTAGGAGCTA (SEQ ID NO: 435)           |
| g15  | GGTTGTGTCACCAGAGTAACAGTCTG (SEQ ID NO: 2311)      |
| g16  | CAGGTTGTGTCACCAGAGTAACAGTC (SEQ ID NO: 2309)      |
| g17  | CACAGGTTGTGTCACCAGAGTAACAG (SEQ ID NO: 2307)      |
| g17+ | CACAGGTTGTGTCACCAGAGTAACAGTCTG (SEQ ID NO: 12312) |
| g18  | ACCACAGGTTGTGTCACCAGAGTAAC (SEQ ID NO: 2305)      |
| g19  | TAACCACAGGTTGTGTCACCAGAGTA (SEQ ID NO: 2303)      |
| g20  | AGTAACCACAGGTTGTGTCACCAGAG (SEQ ID NO: 2301)      |
| g21  | TGTCCAAGCCCGGTTGAAATCTGCCA (SEQ ID NO: 2214)      |
| g22  | GGTAAGTTCTGTCCAAGCCCGGTTGA (SEQ ID NO: 2205)      |
| g23  | CGGTTGAAATCTGCCAGAGCAGGTAC (SEQ ID NO: 2224)      |
| g24  | CCAAGCCCGGTTGAAATCTGCCAGAG (SEQ ID NO: 2217)      |
| g25  | TTCTGTCCAAGCCCGGTTGAAATCTG (SEQ ID NO: 2211)      |
| g26  | TAAGTTCTGTCCAAGCCCGGTTGAAA (SEQ ID NO: 2207)      |
| g27  | AAGGTCACCCACCATCACCTCTGTG (SEQ ID NO: 2122)       |
| g28  | GGTCACCCACCATCACCTCTGTGAT (SEQ ID NO: 2124)       |

|      |                                                                   |
|------|-------------------------------------------------------------------|
| g29  | TCAAGGTCACCCACCATCACCTCTG (SEQ ID NO: 2120)                       |
| g30  | CTGTGATTTTataaCTTGATCAAGCA (SEQ ID NO: 12272)                     |
| g31  | GTGATTTTataaCTTGATCAAGCAGA (SEQ ID NO: 12273)                     |
| g32  | CCACCATCACCTCTGTGATTTTata (SEQ ID NO: 12274)                      |
| g33  | ATTAGAAACACAAGCTAAAGAGCCAA (SEQ ID NO: 8637)                      |
| g34  | TCACCAGAGTAACAGTCTGAGTAGGA (SEQ ID NO: 2318)                      |
| g35  | CACCAGAGTAACAGTCTGAGTAGGAG (SEQ ID NO: 2319)                      |
| g36  | TGTCACCAGAGTAACAGTCTGAGTAG (SEQ ID NO: 2316)                      |
| g37  | GGAAGATGGCATTCTAGTTTGGAGA (SEQ ID NO: 2262)                       |
| g38  | TACCTCCAACATCAAGGAAGATGGCA (SEQ ID NO: 2247)                      |
| g38+ | TACCTCCAACATCAAGGAAGATGGCATTTC (SEQ ID NO: 12285)                 |
| g39  | ATGGCATTCTAGTTTGGAGATGGCA (SEQ ID NO: 2267)                       |
| g40  | ATCTGCCAGAGCAGGTACCTCCAACA (SEQ ID NO: 2232)                      |
| g41  | AGTAACCACAGGTTGTGTCACCAGAGTAAC (SEQ ID NO: 12275)                 |
| g42  | GTAACCACAGGTTGTGTCACCAGAGTAACA (SEQ ID NO: 12276)                 |
| g43  | CCAGAGCAGGTACCTCCAACATCAAGGAAGATGGCATTCTAGTT (SEQ ID NO: 12288)   |
| g45  | CCTTAGTAACCACAGGTTGTGTCACCAGAGTAACAGTCTGAGTAG (SEQ ID NO: 12289)  |
| g46  | GCAGTTTCCTTAGTAACCACAGGTTGTGTCACCAGAGTAACAGTC (SEQ ID NO: 12290)  |
| g47  | GTTTCCTTAGTAACCACAGGTTGTGTCACCAGAGTAACAGTCTGA (SEQ ID NO: 12291)  |
| g48  | CAGAGTAACAGTCTGAGTAGGAGGctaaaatatttgggttttgc (SEQ ID NO: 12292)   |
| g49  | TTGTGTCACCAGAGTAACAGTCTGAGTAGGAGGctaaaatatttg (SEQ ID NO: 12293)  |
| g50  | ACAGGTTGTGTCACCAGAGTAACAGTCTGAGTAGGAGGctaaaata (SEQ ID NO: 12294) |
| g51  | ACAGTCTGAGTAGGAGGctaaaatatttgggttttgcataaagg (SEQ ID NO: 12295)   |

|     |                                                                  |
|-----|------------------------------------------------------------------|
| g52 | ttatcatttttctcatcacCTTCTGCTTGATGATCATCTCGTTGA (SEQ ID NO: 12296) |
| g53 | ccaacttttatcatttttctcatcacCTTCTGCTTGATGATCATC (SEQ ID NO: 12297) |
| g54 | ctgccaacttttatcatttttctcatcacCTTCTGCTTGATGATC (SEQ ID NO: 12298) |
| g55 | gaaaaactctgccaacttttatcatttttctcatcacCTTCTGC (SEQ ID NO: 12299)  |
| g56 | catcacCTTCTGCTTGATGATCATCTCG (SEQ ID NO: 12300)                  |
| g57 | CACCAGAGTAACAGTCTGAGTAGGAGctaaa (SEQ ID NO: 12301)               |
| g58 | GAGTAACAGTCTGAGTAGGAGc (SEQ ID NO: 12302)                        |
| g59 | CCAGAGTAACAGTCTGAGTAGGAGc (SEQ ID NO: 12303)                     |
| g60 | CAGTCTGAGTAGGAGctaaa (SEQ ID NO: 12304)                          |
| g61 | CCAGAGTAACAGTCTGAGTAGGAG (SEQ ID NO: 12305)                      |
| g62 | cCTTCTGCTTGATGATCATCTCG (SEQ ID NO: 12306)                       |

**[0129]** Exemplary gRNA spacer sequences of the disclosure specifically bind to a MAPT pre-mRNA. In some embodiments, the sequence encoding for MAPT comprises a mutation. In some embodiments, the MAPT pre-mRNA is human. In some embodiments, the gRNA spacer sequences of the disclosure specifically bind exon 10 of human MAPT pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind the ESE motif of exon 10 of human MAPT pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind intronic sequences 5' and 3' of exon 10 of human MAPT pre-mRNA. In some aspects, MAPT pre-mRNA gRNA spacer sequences of the disclosure are selected from those listed in Table 4.

**Table 4: Human MAPT exon 10 gRNA spacer sequences**

|    |                                              |
|----|----------------------------------------------|
| g1 | aattatctgcaccttggtagccaga (SEQ ID NO: 438)   |
| g2 | taattatctgcaccttggtagccag (SEQ ID NO: 439)   |
| g3 | tgtgaaggctactcacactgccgcctc (SEQ ID NO: 440) |

**[0130]** Exemplary gRNA spacer sequences of the disclosure specifically bind to a SOD1 pre-mRNA. In some embodiments, the sequence encoding for SOD1 comprises a mutation. In some embodiments, the SOD1 pre-mRNA is human. In some embodiments, gRNA spacer sequences of the disclosure specifically bind at least one of exon 2 and exon 3 of human SOD1 pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind the ESE motif of exon 2 or exon 3 of human SOD1 pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind intronic sequences 5' and 3' of exon 2 or exon 3 of human SOD1 pre-mRNA. In some aspects, SOD1 pre-mRNA gRNA spacer sequences of the disclosure are selected from those listed in Table 5.

**Table 5: Human SOD1 exon 2 and exon 3 gRNA spacer sequences**

| Guide | Spacer Sequence                             |
|-------|---------------------------------------------|
| g1    | TTCATCCTTTGGCCACCGTGTTC (SEQ ID NO: 441)    |
| g2    | TTCCCCACACCTTCACTGGTCCATTA (SEQ ID NO: 442) |
| g3    | TTCAGTCAGTCCTTTAATGCTTCCCC (SEQ ID NO: 443) |
| g4    | GCTGTATTATCTCCAAACTCATGAAC (SEQ ID NO: 444) |

**[0131]** Exemplary gRNA spacer sequences of the disclosure specifically bind to a LRRK2 pre-mRNA. In some embodiments, the sequence encoding for LRRK2 comprises a mutation. In some embodiments, the LRRK2 pre-mRNA is human. In some embodiments, the gRNA spacer sequences of the disclosure specifically bind at least one of exon 2 and exon 3 of human LRRK2 pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind the ESE motif of exon 2 or exon 3 of human LRRK2 pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind intronic sequences 5' and 3' of exon 2 or exon 3 of human LRRK2 pre-mRNA. In some aspects, LRRK2 pre-mRNA gRNA spacer sequences of the disclosure are selected from those listed in Table 6.

**Table 6: Human LRRK2 exon 2 and exon 3 gRNA spacer sequences**

| Guide | Spacer Sequence                             |
|-------|---------------------------------------------|
| g1    | ctctcatataggagtccaagacgatc (SEQ ID NO: 446) |

|    |                                              |
|----|----------------------------------------------|
| g2 | ccaacatcctgggggtcccattaagct (SEQ ID NO: 447) |
|----|----------------------------------------------|

**[0132]** Exemplary gRNA spacer sequences of the disclosure specifically bind to a SMN2 pre-mRNA. In some embodiments, the sequence encoding for SMN2 comprises a mutation. In some embodiments, the SMN2 pre-mRNA is human. In some embodiments, gRNA spacer sequences of the disclosure specifically bind an intronic region of the SMN2 pre-mRNA. In some embodiments, the gRNA spacer sequences of the disclosure bind the ISS motif of an intron of SMN2. In some embodiments, SMN2-targeting gRNA spacer sequences of the disclosure promote inclusion of an exon that is excluded during pre-mRNA splicing due to the presence of a mutation or SNP. In some embodiments, the ISS motif occurs between exon 7 and exon 8. In some aspects, SMN2 pre-mRNA gRNA spacer sequences of the disclosure are selected from those listed in Table B.

**Table B: Human SMN2 gRNA spacer sequences**

| Guide | Spacer Sequence                                    |
|-------|----------------------------------------------------|
| g1    | AAGATTCACCTTTCATAATGCTGGCAGACTT (SEQ ID NO: 12317) |
| g2    | AGATTCACCTTTCATAATGCTGGCAGACTTA (SEQ ID NO: 12318) |
| g3    | GTAAGATTCACCTTTCATAATGCTGGCAGAC (SEQ ID NO: 12319) |
| g4    | GATTCACCTTTCATAATGCTGGCAGAC (SEQ ID NO: 12320)     |
| g5    | ATTCACCTTTCATAATGCTGGCAGACT (SEQ ID NO: 12321)     |

### **RNA-guided RNA-binding Proteins**

**[0133]** In some embodiments of the compositions of the disclosure, gRNAs correspond to target RNA molecules and an RNA-guided RNA binding protein. In some embodiments, the gRNAs correspond to an RNA-guided RNA binding fusion protein, wherein the fusion protein comprises first and second RNA binding proteins. In some embodiments, the RNA-binding protein is a deactivated RNA-binding protein, e.g., a deactivated Cas or catalytic dead Cas protein. In some embodiments, the first RNA-binding protein in the fusion protein is a deactivated RNA-binding protein, e.g., a deactivated Cas or catalytic dead Cas protein. In some embodiments, along a sequence encoding the RNA-binding fusion protein, the

sequence encoding the first RNA binding protein is positioned 5' of the sequence encoding the second RNA binding protein. In some embodiments, along a sequence encoding the fusion protein, the sequence encoding the first RNA binding protein is positioned 3' of the sequence encoding the second RNA binding protein.

**[0134]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA binding protein comprises a sequence isolated or derived from a protein capable of binding an RNA molecule. In some embodiments, the sequence encoding the RNA binding protein comprises a sequence isolated or derived from a protein capable of selectively binding an RNA molecule and not binding a DNA molecule, a mammalian DNA molecule or any DNA molecule. In some embodiments, the sequence encoding the RNA binding protein comprises a sequence isolated or derived from a protein capable of binding an RNA molecule and inducing a break in the RNA molecule. In some embodiments, the sequence encoding the RNA binding protein comprises a sequence isolated or derived from a protein capable of binding an RNA molecule, inducing a break in the RNA molecule, and not binding a DNA molecule, a mammalian DNA molecule or any DNA molecule. In some embodiments, the sequence encoding the RNA binding protein comprises a sequence isolated or derived from a protein capable of binding an RNA molecule, inducing a break in the RNA molecule, and neither binding nor inducing a break in a DNA molecule, a mammalian DNA molecule or any DNA molecule.

**[0135]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA-guided RNA binding protein comprises a sequence isolated or derived from a protein with no DNA nuclease activity.

**[0136]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA-guided RNA binding protein disclosed herein comprises a sequence isolated or derived from a CRISPR Cas protein. In some embodiments, the CRISPR Cas protein is not a Type II CRISPR Cas protein. In some embodiments, the CRISPR Cas protein is not a Cas9 protein.

**[0137]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA-guided RNA binding protein comprises a Type VI CRISPR Cas protein or portion thereof. In some embodiments, the Type VI CRISPR Cas protein comprises a Cas13 protein or portion thereof. Exemplary Cas13 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, a bacteria or an archaea. Exemplary Cas13 proteins of the disclosure may be isolated or derived from any species, including, but not limited to, *Leptotrichia wadei*, *Listeria seeligeri* serovar 1/2b (strain ATCC 35967 / DSM

20751 / CIP 100100 / SLCC 3954), *Lachnospiraceae* bacterium, *Clostridium aminophilum* DSM 10710, *Carnobacterium gallinarum* DSM 4847, *Paludibacter propionigenes* WB4, *Listeria weihenstephanensis* FSL R9-0317, *Listeria weihenstephanensis* FSL R9-0317, bacterium FSL M6-0635 (*Listeria newyorkensis*), *Leptotrichia wadei* F0279, *Rhodobacter capsulatus* SB 1003, *Rhodobacter capsulatus* R121, *Rhodobacter capsulatus* DE442 and *Corynebacterium ulcerans*. Exemplary Cas13 proteins of the disclosure may be DNA nuclease inactivated. Exemplary Cas13 proteins of the disclosure include, but are not limited to, Cas13a, Cas13b, Cas13c, Cas13d and orthologs thereof. Exemplary Cas13b proteins of the disclosure include, but are not limited to, subtypes 1 and 2 referred to herein as Csx27 and Csx28, respectively.

[0138] Exemplary Cas13a proteins include, but are not limited to:

| Cas13a number | Cas13a abbreviation | Organism name                       | Accession number | Direct Repeat sequence                                 |
|---------------|---------------------|-------------------------------------|------------------|--------------------------------------------------------|
| Cas13a1       | LshCas13a           | Leptotrichia shahii                 | WP_018451595.1   | CCACCCCAATATCGAAGGGG ACTAAAAC (SEQ ID NO: 323)         |
| Cas13a2       | LwaCas13a           | Leptotrichia wadei                  | WP_021746774.1   | GATTTAGACTACCCCAAAAAC GAAGGGGACTAAAAC (SEQ ID NO: 324) |
| Cas13a3       | LseCas13a           | Listeria seeligeri                  | WP_012985477.1   | GTAAGAGACTACCTCTATATG AAAGAGGACTAAAAC (SEQ ID NO: 325) |
| Cas13a4       | LbmCas13a           | Lachnospiraceae bacterium MA2020    | WP_044921188.1   | GTATTGAGAAAAGCCAGATA TAGTTGGCAATAGAC (SEQ ID NO: 326)  |
| Cas13a5       | LbnCas13a           | Lachnospiraceae bacterium NK4A179   | WP_022785443.1   | GTTGATGAGAAGAGCCCAAG ATAGAGGGCAATAAC (SEQ ID NO: 327)  |
| Cas13a6       | CamCas13a           | [Clostridium] aminophilum DSM 10710 | WP_031473346.1   | GTCTATTGCCCTCTATATCGG GCTGTTCTCCAAAC (SEQ ID NO: 328)  |
| Cas13a7       | CgaCas13a           | Carnobacterium gallinarum DSM 4847  | WP_034560163.1   | ATTAAAGACTACCTCTAAATG TAAGAGGACTATAAC (SEQ ID NO: 329) |
| Cas13a8       | Cga2Cas13a          | Carnobacterium gallinarum DSM 4847  | WP_034563842.1   | AATATAAACTACCTCTAAATG TAAGAGGACTATAAC (SEQ ID NO: 330) |

|          |            |                                                            |                |                                                        |
|----------|------------|------------------------------------------------------------|----------------|--------------------------------------------------------|
| Cas13a9  | Pprcas13a  | Paludibacter propionigenes WB4                             | WP_013443710.1 | CTTGTGGATTATCCCAAATTGAAGGGAACACTACAAC (SEQ ID NO: 331) |
| Cas13a10 | LweCas13a  | Listeria weihenstephanensis FSL R9-0317                    | WP_036059185.1 | GATTTAGAGTACCTCAAAATAGAAGAGGTCTAAAAC (SEQ ID NO: 332)  |
| Cas13a11 | LbfCas13a  | Listeriaceae bacterium FSL M6-0635 (Listeria newyorkensis) | WP_036091002.1 | GATTTAGAGTACCTCAAAACAAAGAGGACTAAAAC (SEQ ID NO: 333)   |
| Cas13a12 | Lwa2cas13a | Leptotrichia wadei F0279                                   | WP_021746774.1 | GATATAGATAACCCCCAAAAACGAAGGGATCTAAAAC (SEQ ID NO: 334) |
| Cas13a13 | RcsCas13a  | Rhodobacter capsulatus SB 1003                             | WP_013067728.1 | GCCTCACATCACCGCCAAGACGACGGCGGACTGAAC (SEQ ID NO: 335)  |
| Cas13a14 | RcrCas13a  | Rhodobacter capsulatus R121                                | WP_023911507.1 | GCCTCACATCACCGCCAAGACGACGGCGGACTGAAC (SEQ ID NO: 336)  |
| Cas13a15 | RcdCas13a  | Rhodobacter capsulatus DE442                               | WP_023911507.1 | GCCTCACATCACCGCCAAGACGACGGCGGACTGAAC (SEQ ID NO: 337)  |

[0139] Exemplary wild type Cas13a proteins of the disclosure may comprise or consist of the amino acid sequence of SEQ ID NO: 338.

[0140] Exemplary Cas13b proteins include, but are not limited to:

| Species                             | Cas13b Accession | Cas13b Size (aa) |
|-------------------------------------|------------------|------------------|
| Paludibacter propionigenes WB4      | WP_013446107.1   | 1155             |
| Prevotella sp. P5-60                | WP_044074780.1   | 1091             |
| Prevotella sp. P4-76                | WP_044072147.1   | 1091             |
| Prevotella sp. P5-125               | WP_044065294.1   | 1091             |
| Prevotella sp. P5-119               | WP_042518169.1   | 1091             |
| Capnocytophaga canimorsus Cc5       | WP_013997271.1   | 1200             |
| Phaeodactylibacter xiamenensis      | WP_044218239.1   | 1132             |
| Porphyromonas gingivalis W83        | WP_005873511.1   | 1136             |
| Porphyromonas gingivalis F0570      | WP_021665475.1   | 1136             |
| Porphyromonas gingivalis ATCC 33277 | WP_012458151.1   | 1136             |
| Porphyromonas gingivalis F0185      | ERJ81987.1       | 1136             |
| Porphyromonas gingivalis F0185      | WP_021677657.1   | 1136             |
| Porphyromonas gingivalis SJD2       | WP_023846767.1   | 1136             |
| Porphyromonas gingivalis F0568      | ERJ65637.1       | 1136             |
| Porphyromonas gingivalis W4087      | ERJ87335.1       | 1136             |
| Porphyromonas gingivalis W4087      | WP_021680012.1   | 1136             |

|                                                        |                |      |
|--------------------------------------------------------|----------------|------|
| <i>Porphyromonas gingivalis</i> F0568                  | WP_021663197.1 | 1136 |
| <i>Porphyromonas gingivalis</i>                        | WP_061156637.1 | 1136 |
| <i>Porphyromonas gulae</i>                             | WP_039445055.1 | 1136 |
| <i>Bacteroides pyogenes</i> F0041                      | ERI81700.1     | 1116 |
| <i>Bacteroides pyogenes</i> JCM 10003                  | WP_034542281.1 | 1116 |
| <i>Alistipes</i> sp. ZOR0009                           | WP_047447901.1 | 954  |
| <i>Flavobacterium branchiophilum</i> FL-15             | WP_014084666.1 | 1151 |
| <i>Prevotella</i> sp. MA2016                           | WP_036929175.1 | 1323 |
| <i>Myroides odoratimimus</i> CCUG 10230                | EHO06562.1     | 1160 |
| <i>Myroides odoratimimus</i> CCUG 3837                 | EKB06014.1     | 1158 |
| <i>Myroides odoratimimus</i> CCUG 3837                 | WP_006265509.1 | 1158 |
| <i>Myroides odoratimimus</i> CCUG 12901                | WP_006261414.1 | 1158 |
| <i>Myroides odoratimimus</i> CCUG 12901                | EHO08761.1     | 1158 |
| <i>Myroides odoratimimus</i> (NZ_CP013690.1)           | WP_058700060.1 | 1160 |
| <i>Bergeyella zoohelcum</i> ATCC 43767                 | EKB54193.1     | 1225 |
| <i>Capnocytophaga cynodegmi</i>                        | WP_041989581.1 | 1219 |
| <i>Bergeyella zoohelcum</i> ATCC 43767                 | WP_002664492.1 | 1225 |
| <i>Flavobacterium</i> sp. 316                          | WP_045968377.1 | 1156 |
| <i>Psychroflexus torquis</i> ATCC 700755               | WP_015024765.1 | 1146 |
| <i>Flavobacterium columnare</i> ATCC 49512             | WP_014165541.1 | 1180 |
| <i>Flavobacterium columnare</i>                        | WP_060381855.1 | 1214 |
| <i>Flavobacterium columnare</i>                        | WP_063744070.1 | 1214 |
| <i>Flavobacterium columnare</i>                        | WP_065213424.1 | 1215 |
| <i>Chryseobacterium</i> sp. YR477                      | WP_047431796.1 | 1146 |
| <i>Riemerella anatipestifer</i> ATCC 11845 = DSM 15868 | WP_004919755.1 | 1096 |
| <i>Riemerella anatipestifer</i> RA-CH-2                | WP_015345620.1 | 949  |
| <i>Riemerella anatipestifer</i>                        | WP_049354263.1 | 949  |
| <i>Riemerella anatipestifer</i>                        | WP_061710138.1 | 951  |
| <i>Riemerella anatipestifer</i>                        | WP_064970887.1 | 1096 |
| <i>Prevotella saccharolytica</i> F0055                 | EKY00089.1     | 1151 |
| <i>Prevotella saccharolytica</i> JCM 17484             | WP_051522484.1 | 1152 |
| <i>Prevotella buccae</i> ATCC 33574                    | EFU31981.1     | 1128 |
| <i>Prevotella buccae</i> ATCC 33574                    | WP_004343973.1 | 1128 |
| <i>Prevotella buccae</i> D17                           | WP_004343581.1 | 1128 |
| <i>Prevotella</i> sp. MSX73                            | WP_007412163.1 | 1128 |
| <i>Prevotella pallens</i> ATCC 700821                  | EGQ18444.1     | 1126 |
| <i>Prevotella pallens</i> ATCC 700821                  | WP_006044833.1 | 1126 |
| <i>Prevotella intermedia</i> ATCC 25611 = DSM 20706    | WP_036860899.1 | 1127 |
| <i>Prevotella intermedia</i>                           | WP_061868553.1 | 1121 |
| <i>Prevotella intermedia</i> 17                        | AFJ07523.1     | 1135 |
| <i>Prevotella intermedia</i>                           | WP_050955369.1 | 1133 |
| <i>Prevotella intermedia</i>                           | BAU18623.1     | 1134 |
| <i>Prevotella intermedia</i> ZT                        | KJJ86756.1     | 1126 |
| <i>Prevotella aurantiaca</i> JCM 15754                 | WP_025000926.1 | 1125 |
| <i>Prevotella pleuritidis</i> F0068                    | WP_021584635.1 | 1140 |
| <i>Prevotella pleuritidis</i> JCM 14110                | WP_036931485.1 | 1117 |

|                                           |                |      |
|-------------------------------------------|----------------|------|
| Prevotella falsenii DSM 22864 = JCM 15124 | WP_036884929.1 | 1134 |
| Porphyromonas gulae                       | WP_039418912.1 | 1176 |
| Porphyromonas sp. COT-052 OH4946          | WP_039428968.1 | 1176 |
| Porphyromonas gulae                       | WP_039442171.1 | 1175 |
| Porphyromonas gulae                       | WP_039431778.1 | 1176 |
| Porphyromonas gulae                       | WP_046201018.1 | 1176 |
| Porphyromonas gulae                       | WP_039434803.1 | 1176 |
| Porphyromonas gulae                       | WP_039419792.1 | 1120 |
| Porphyromonas gulae                       | WP_039426176.1 | 1120 |
| Porphyromonas gulae                       | WP_039437199.1 | 1120 |
| Porphyromonas gingivalis TDC60            | WP_013816155.1 | 1120 |
| Porphyromonas gingivalis ATCC 33277       | WP_012458414.1 | 1120 |
| Porphyromonas gingivalis A7A1-28          | WP_058019250.1 | 1176 |
| Porphyromonas gingivalis JCVI SC001       | EOA10535.1     | 1176 |
| Porphyromonas gingivalis W50              | WP_005874195.1 | 1176 |
| Porphyromonas gingivalis                  | WP_052912312.1 | 1176 |
| Porphyromonas gingivalis AJW4             | WP_053444417.1 | 1120 |
| Porphyromonas gingivalis                  | WP_039417390.1 | 1120 |
| Porphyromonas gingivalis                  | WP_061156470.1 | 1120 |

**[0141]** Exemplary wild type *Bergeyella zoohelcum* ATCC 43767 Cas13b (BzCas13b) proteins of the disclosure may comprise or consist of the amino acid sequence of SEQ ID NO: 339.

**[0142]** In some embodiments of the compositions of the disclosure, the sequence encoding the RNA binding protein comprises a sequence isolated or derived from a Cas13d protein. Cas13d is an effector of the type VI-D CRISPR-Cas systems. In some embodiments, the Cas13d protein is an RNA-guided RNA endonuclease enzyme that can cut or bind RNA. In some embodiments, the Cas13d protein can include one or more higher eukaryotes and prokaryotes nucleotide-binding (HEPN) domains. In some embodiments, the Cas13d protein can include either a wild-type or mutated HEPN domain. In some embodiments, the Cas13d protein includes a mutated HEPN domain that cannot cut RNA but can process guide RNA. In some embodiments, the Cas13d protein does not require a protospacer flanking sequence. Also see WO Publication No. WO2019/040664 & US2019/0062724, which is incorporated herein by reference in its entirety, for further examples and sequences of Cas13d protein, without limitation.

**[0143]** In some embodiments, Cas13d sequences of the disclosure include without limitation SEQ ID NOS: 1-296 of WO 2019/040664, so numbered herein and included herewith.

[0144] **SEQ ID NO: 1** is an exemplary Cas13d sequence from *Eubacterium siraeum* containing a HEPN site.

[0145] **SEQ ID NO: 2** is an exemplary Cas13d sequence from *Eubacterium siraeum* containing a mutated HEPN site.

[0146] **SEQ ID NO: 3** is an exemplary Cas13d sequence from uncultured *Ruminococcus* sp. containing a HEPN site.

[0147] **SEQ ID NO: 4** is an exemplary Cas13d sequence from uncultured *Ruminococcus* sp. containing a mutated HEPN site.

[0148] **SEQ ID NO: 5** is an exemplary Cas13d sequence from Gut\_metagenome\_contig2791000549.

[0149] **SEQ ID NO: 6** is an exemplary Cas13d sequence from Gut\_metagenome\_contig855000317

[0150] **SEQ ID NO: 7** is an exemplary Cas13d sequence from Gut\_metagenome\_contig3389000027.

[0151] **SEQ ID NO: 8** is an exemplary Cas13d sequence from Gut\_metagenome\_contig8061000170.

[0152] **SEQ ID NO: 9** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1509000299.

[0153] **SEQ ID NO: 10** is an exemplary Cas13d sequence from Gut\_metagenome\_contig9549000591.

[0154] **SEQ ID NO: 11** is an exemplary Cas13d sequence from Gut\_metagenome\_contig71000500.

[0155] **SEQ ID NO: 12** is an exemplary Cas13d sequence from human gut metagenome.

[0156] **SEQ ID NO: 13** is an exemplary Cas13d sequence from Gut\_metagenome\_contig3915000357.

[0157] **SEQ ID NO: 14** is an exemplary Cas13d sequence from Gut\_metagenome\_contig4719000173.

[0158] **SEQ ID NO: 15** is an exemplary Cas13d sequence from Gut\_metagenome\_contig6929000468.

**[0159] SEQ ID NO: 16** is an exemplary Cas13d sequence from Gut\_metagenome\_contig7367000486.

**[0160] SEQ ID NO: 17** is an exemplary Cas13d sequence from Gut\_metagenome\_contig7930000403.

**[0161] SEQ ID NO: 18** is an exemplary Cas13d sequence from Gut\_metagenome\_contig993000527.

**[0162] SEQ ID NO: 19** is an exemplary Cas13d sequence from Gut\_metagenome\_contig6552000639.

**[0163] SEQ ID NO: 20** is an exemplary Cas13d sequence from Gut\_metagenome\_contigl932000246.

**[0164] SEQ ID NO: 21** is an exemplary Cas13d sequence from Gut\_metagenome\_contigl2963000286.

**[0165] SEQ ID NO: 22** is an exemplary Cas13d sequence from Gut\_metagenome\_contig2952000470.

**[0166] SEQ ID NO: 23** is an exemplary Cas13d sequence from Gut\_metagenome\_contig451000394.

**[0167] SEQ ID NO: 24** is an exemplary Cas13d sequence from Eubacterium\_siraeum\_DSM\_15702.

**[0168] SEQ ID NO: 25** is an exemplary Cas13d sequence from gut\_metagenome\_P19E0k2120140920,\_c369000003.

**[0169] SEQ ID NO: 26** is an exemplary Cas13d sequence from Gut\_metagenome\_contig7593000362.

**[0170] SEQ ID NO: 27** is an exemplary Cas13d sequence from Gut\_metagenome\_contigl2619000055.

**[0171] SEQ ID NO: 28** is an exemplary Cas13d sequence from Gut\_metagenome\_contigl405000151.

**[0172] SEQ ID NO: 29** is an exemplary Cas13d sequence from Chicken\_gut\_metagenome\_c298474.

**[0173] SEQ ID NO: 30** is an exemplary Cas13d sequence from Gut\_metagenome\_contigl516000227.

**[0174] SEQ ID NO: 31** is an exemplary Cas13d sequence from Gut\_metagenome\_contigl838000319.

[0175] **SEQ ID NO: 32** is an exemplary Cas13d sequence from Gut\_metagenome\_contig13123000268.

[0176] **SEQ ID NO: 33** is an exemplary Cas13d sequence from Gut\_metagenome\_contig5294000434.

[0177] **SEQ ID NO: 34** is an exemplary Cas13d sequence from Gut\_metagenome\_contig6415000192.

[0178] **SEQ ID NO: 35** is an exemplary Cas13d sequence from Gut\_metagenome\_contig6144000300.

[0179] **SEQ ID NO: 36** is an exemplary Cas13d sequence from Gut\_metagenome\_contig9118000041.

[0180] **SEQ ID NO: 37** is an exemplary Cas13d sequence from Activated\_sludge\_metagenome\_transcript\_124486.

[0181] **SEQ ID NO: 38** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1322000437.

[0182] **SEQ ID NO: 39** is an exemplary Cas13d sequence from Gut\_metagenome\_contig4582000531.

[0183] **SEQ ID NO: 40** is an exemplary Cas13d sequence from Gut\_metagenome\_contig9190000283.

[0184] **SEQ ID NO: 41** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1709000510.

[0185] **SEQ ID NO: 42** is an exemplary Cas13d sequence from M24\_(LSQX01212483\_Anaerobic\_digester\_metagenome) with a HEPN domain.

[0186] **SEQ ID NO: 43** is an exemplary Cas13d sequence from Gut\_metagenome\_contig3833000494.

[0187] **SEQ ID NO: 44** is an exemplary Cas13d sequence from Activated\_sludge\_metagenome\_transcript\_117355.

[0188] **SEQ ID NO: 45** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1061000330.

[0189] **SEQ ID NO: 46** is an exemplary Cas13d sequence from Gut\_metagenome\_contig338000322 from sheep gut metagenome.

[0190] **SEQ ID NO: 47** is an exemplary Cas13d sequence from human gut metagenome.

[0191] **SEQ ID NO: 48** is an exemplary Cas13d sequence from Gut\_metagenome\_contig9530000097.

[0192] **SEQ ID NO: 49** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1750000258.

[0193] **SEQ ID NO: 50** is an exemplary Cas13d sequence from Gut\_metagenome\_contig5377000274.

[0194] **SEQ ID NO: 51** is an exemplary Cas13d sequence from gut\_metagenome\_P19E0k2120140920\_c248000089.

[0195] **SEQ ID NO: 52** is an exemplary Cas13d sequence from Gut\_metagenome\_contig11400000031.

[0196] **SEQ ID NO: 53** is an exemplary Cas13d sequence from Gut\_metagenome\_contig7940000191.

[0197] **SEQ ID NO: 54** is an exemplary Cas13d sequence from Gut\_metagenome\_contig6049000251.

[0198] **SEQ ID NO: 55** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1137000500.

[0199] **SEQ ID NO: 56** is an exemplary Cas13d sequence from Gut\_metagenome\_contig9368000105.

[0200] **SEQ ID NO: 57** is an exemplary Cas13d sequence from Gut\_metagenome\_contig546000275.

[0201] **SEQ ID NO: 58** is an exemplary Cas13d sequence from Gut\_metagenome\_contig7216000573.

[0202] **SEQ ID NO: 59** is an exemplary Cas13d sequence from Gut\_metagenome\_contig4806000409.

[0203] **SEQ ID NO: 60** is an exemplary Cas13d sequence from Gut\_metagenome\_contig10762000480.

[0204] **SEQ ID NO: 61** is an exemplary Cas13d sequence from Gut\_metagenome\_contig4114000374.

[0205] **SEQ ID NO: 62** is an exemplary Cas13d sequence from *Ruminococcus\_flavefaciens*\_FD1.

[0206] **SEQ ID NO: 63** is an exemplary Cas13d sequence from Gut\_metagenome\_contig7093000170.

[0207] **SEQ ID NO: 64** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1113000384.

[0208] **SEQ ID NO: 65** is an exemplary Cas13d sequence from Gut\_metagenome\_contig6403000259.

[0209] **SEQ ID NO: 66** is an exemplary Cas13d sequence from Gut\_metagenome\_contig6193000124.

[0210] **SEQ ID NO: 67** is an exemplary Cas13d sequence from Gut\_metagenome\_contig721000619.

[0211] **SEQ ID NO: 68** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1666000270.

[0212] **SEQ ID NO: 69** is an exemplary Cas13d sequence from Gut\_metagenome\_contig2002000411.

[0213] **SEQ ID NO: 70** is an exemplary Cas13d sequence from *Ruminococcus\_albus*.

[0214] **SEQ ID NO: 71** is an exemplary Cas13d sequence from Gut\_metagenome\_contig13552000311.

[0215] **SEQ ID NO: 72** is an exemplary Cas13d sequence from Gut\_metagenome\_contig10037000527.

[0216] **SEQ ID NO: 73** is an exemplary Cas13d sequence from Gut\_metagenome\_contig238000329.

[0217] **SEQ ID NO: 74** is an exemplary Cas13d sequence from Gut\_metagenome\_contig2643000492.

[0218] **SEQ ID NO: 75** is an exemplary Cas13d sequence from Gut\_metagenome\_contig874000057.

[0219] **SEQ ID NO: 76** is an exemplary Cas13d sequence from Gut\_metagenome\_contig4781000489.

[0220] **SEQ ID NO: 77** is an exemplary Cas13d sequence from Gut\_metagenome\_contig12144000352.

[0221] **SEQ ID NO: 78** is an exemplary Cas13d sequence from Gut\_metagenome\_contig5590000448.

[0222] **SEQ ID NO: 79** is an exemplary Cas13d sequence from Gut\_metagenome\_contig9269000031.

[0223] **SEQ ID NO: 80** is an exemplary Cas13d sequence from Gut\_metagenome\_contig8537000520.

[0224] **SEQ ID NO: 81** is an exemplary Cas13d sequence from Gut\_metagenome\_contig1845000130.

[0225] **SEQ ID NO: 82** is an exemplary Cas13d sequence from gut\_metagenome\_P13E0k2l20140920\_c3000072.

[0226] **SEQ ID NO: 83** is an exemplary Cas13d sequence from gut\_metagenome\_P1E0k2l20140920\_cI000078.

[0227] **SEQ ID NO: 84** is an exemplary Cas13d sequence from Gut\_metagenome\_contig12990000099.

[0228] **SEQ ID NO: 85** is an exemplary Cas13d sequence from Gut\_metagenome\_contig525000349.

[0229] **SEQ ID NO: 86** is an exemplary Cas13d sequence from Gut\_metagenome\_contig7229000302.

[0230] **SEQ ID NO: 87** is an exemplary Cas13d sequence from Gut\_metagenome\_contig3227000343.

[0231] **SEQ ID NO: 88** is an exemplary Cas13d sequence from Gut\_metagenome\_contig7030000469.

[0232] **SEQ ID NO: 89** is an exemplary Cas13d sequence from Gut\_metagenome\_contig5149000068.

[0233] **SEQ ID NO: 90** is an exemplary Cas13d sequence from Gut\_metagenome\_contig400200045.

[0234] **SEQ ID NO: 91** is an exemplary Cas13d sequence from Gut\_metagenome\_contig10420000446.

[0235] **SEQ ID NO: 92** is an exemplary Cas13d sequence from new\_flavefaciens\_strain\_XPD3002 (CasRx).

[0236] **SEQ ID NO: 93** is an exemplary Cas13d sequence from M26\_Gut\_metagenome\_contig698000307.

- [0237] **SEQ ID NO: 94** is an exemplary Cas13d sequence from M36\_Uncultured\_*Eubacterium\_sp\_TS28\_c40956*.
- [0238] **SEQ ID NO: 95** is an exemplary Cas13d sequence from M12\_gut\_metagenome\_P25C0k2l20140920\_c134000066.
- [0239] **SEQ ID NO: 96** is an exemplary Cas13d sequence from human gut metagenome.
- [0240] **SEQ ID NO: 97** is an exemplary Cas13d sequence from M10\_gut\_metagenome\_P25C90k2l20140920\_c28000041.
- [0241] **SEQ ID NO: 98** is an exemplary Cas13d sequence from 30 M1I\_gut\_metagenome\_P25C7k2l20140920\_c4078000105.
- [0242] **SEQ ID NO: 99** is an exemplary Cas13d sequence from gut\_metagenome\_P25C0k2l20140920\_c32000045.
- [0243] **SEQ ID NO: 100** is an exemplary Cas13d sequence from M13\_gut\_metagenome\_P23C7k2l20140920\_c3000067.
- [0244] **SEQ ID NO: 101** is an exemplary Cas13d sequence from M5\_gut\_metagenome\_P18E90k2l20140920.
- [0245] **SEQ ID NO: 102** is an exemplary Cas13d sequence from M2l\_gut\_metagenome\_P18E0k2l20140920.
- [0246] **SEQ ID NO: 103** is an exemplary Cas13d sequence from M7\_gut\_metagenome\_P38C7k2l20140920\_c4841000003.
- [0247] **SEQ ID NO: 104** is an exemplary Cas13d sequence from *Ruminococcus\_bicirculans*.
- [0248] **SEQ ID NO: 105** is an exemplary Cas13d sequence.
- [0249] **SEQ ID NO: 106** is an exemplary Cas13d consensus sequence.
- [0250] **SEQ ID NO: 107** is an exemplary Cas13d sequence from M18\_gut\_metagenome\_P22E0k2l20140920\_c3395000078.
- [0251] **SEQ ID NO: 108** is an exemplary Cas13d sequence from M17\_gut\_metagenome\_P22E90k2l20140920\_c114.
- [0252] **SEQ ID NO: 109** is an exemplary Cas13d sequence from *Ruminococcus\_sp\_CAG57*.
- [0253] **SEQ ID NO: 110** is an exemplary Cas13d sequence from gut\_metagenome\_P11E90k2l20140920\_c43000123.

- [0254] **SEQ ID NO: 111** is an exemplary Cas13d sequence from M6\_gut\_metagenome\_P13E90k2120\_1\_40920\_c7000009.
- [0255] **SEQ ID NO: 112** is an exemplary Cas13d sequence from M19\_gut\_metagenome\_P17E90k2120140920.
- [0256] **SEQ ID NO: 113** is an exemplary Cas13d sequence from gut\_metagenome\_P17E0k2120140920\_c87000043.
- [0257] **SEQ ID NO: 114** is an exemplary human codon optimized *Eubacterium siraeum* Cas13d nucleic acid sequence.
- [0258] **SEQ ID NO: 115** is an exemplary human codon optimized *Eubacterium siraeum* Cas13d nucleic acid sequence with a mutant HEPN domain.
- [0259] **SEQ ID NO: 116** is an exemplary human codon-optimized *Eubacterium siraeum* Cas13d nucleic acid sequence with N-terminal NLS.
- [0260] **SEQ ID NO: 117** is an exemplary human codon-optimized *Eubacterium siraeum* Cas13d nucleic acid sequence with N- and C-terminal NLS tags.
- [0261] **SEQ ID NO: 118** is an exemplary human codon-optimized uncultured *Ruminococcus* sp. Cas13d 30 nucleic acid sequence.
- [0262] **SEQ ID NO: 119** is an exemplary human codon-optimized uncultured *Ruminococcus* sp. Cas13d nucleic acid sequence with a mutant HEPN domain.
- [0263] **SEQ ID NO: 120** is an exemplary human codon-optimized uncultured *Ruminococcus* sp. Cas13d nucleic acid sequence with N-terminal NLS.
- [0264] **SEQ ID NO: 121** is an exemplary human codon-optimized uncultured *Ruminococcus* sp. Cas13d nucleic acid sequence with N- and C-terminal NLS tags.
- [0265] **SEQ ID NO: 122** is an exemplary human codon-optimized uncultured *Ruminococcus flavefaciens* FDI Cas13d nucleic acid sequence.
- [0266] **SEQ ID NO: 123** is an exemplary human codon-optimized uncultured *Ruminococcus flavefaciens* FDI Cas13d nucleic acid sequence with mutated HEPN domain.
- [0267] **SEQ ID NO: 124** is an exemplary Cas13d nucleic acid sequence from *Ruminococcus bicirculans*.
- [0268] **SEQ ID NO: 125** is an exemplary Cas13d nucleic acid sequence from *Eubacterium siraeum*.
- [0269] **SEQ ID NO: 126** is an exemplary Cas13d nucleic acid sequence from *Ruminococcus flavefaciens* FD1.

[0270] **SEQ ID NO: 127** is an exemplary Cas13d nucleic acid sequence from *Ruminococcus albus*.

[0271] **SEQ ID NO: 128** is an exemplary Cas13d nucleic acid sequence from *Ruminococcus flavefaciens* XPD.

[0272] **SEQ ID NO: 129** is an exemplary consensus DR nucleic acid sequence for *E. siraeum* Cas13d.

[0273] **SEQ ID NO: 130** is an exemplary consensus DR nucleic acid sequence for *Rum. Sp.* Cas13d.

[0274] **SEQ ID NO: 131** is an exemplary consensus DR nucleic acid sequence for *Rum. Flavefaciens strain XPD3002* Cas13d (CasRx).

[0275] **SEQ ID NOS: 132-137** are exemplary consensus DR nucleic acid sequences.

[0276] **SEQ ID NO: 138** is an exemplary 50% consensus sequence for seven full-length Cas13d orthologues.

[0277] **SEQ ID NO: 139** is an exemplary Cas13d nucleic acid sequence from Gut metagenome PIEO.

[0278] **SEQ ID NO: 140** is an exemplary Cas13d nucleic acid sequence from *Anaerobic digester*.

[0279] **SEQ ID NO: 141** is an exemplary Cas13d nucleic acid sequence from *Ruminococcus sp. CAG:57*.

[0280] **SEQ ID NO: 142** is an exemplary human codon-optimized uncultured Gut metagenome PIEO Cas13d nucleic acid sequence.

[0281] **SEQ ID NO: 143** is an exemplary human codon-optimized Anaerobic Digester Cas13d nucleic acid sequence.

[0282] **SEQ ID NO: 144** is an exemplary human codon-optimized *Ruminococcus flavefaciens* XPD Cas13d nucleic acid sequence.

[0283] **SEQ ID NO: 145** is an exemplary human codon-optimized *Ruminococcus albus* Cas13d nucleic acid sequence.

[0284] **SEQ ID NO: 146** is an exemplary processing of the *Ruminococcus sp. CAG:57* CRISPR array.

[0285] **SEQ ID NO: 147** is an exemplary Cas13d protein sequence from contig emb |OBVH01003037.1, human gut metagenome sequence (also found in WGS contigs emb |OBXZ01000094.1 | and emb |OBJFO1000033.1.

[0286] **SEQ ID NO: 148** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO:147).

[0287] **SEQ ID NO: 149** is an exemplary Cas13d protein sequence from contig tpg |DBYI01000091.1 | (Uncultivated *Ruminococcus flavefaciens* UBA1190 assembled from bovine gut metagenome).

[0288] **SEQ ID NOS: 150-152** are exemplary consensus DR nucleic acid sequences (goes with SEQ ID NO: 149).

[0289] **SEQ ID NO: 153** is an exemplary Cas13d protein sequence from contig tpg |DJXD01000002.1 | (uncultivated *Ruminococcus* assembly, UBA7013, from sheep gutmetagenome).

[0290] **SEQ ID NO: 154** is an exemplary consensus DRnucleic acid sequence (goes with SEQ ID NO: 153).

[0291] **SEQ ID NO: 155** is an exemplary Cas13d protein sequence from contig OGZC01000639.1 (human gut metagenome assembly).

[0292] **SEQ ID NOS: 156-177** are exemplary consensus DR nucleic acid sequences (goes with SEQ ID NO: 155).

[0293] **SEQ ID NO: 158** is an exemplary Cas13d protein sequence from contig emb |OHBM01000764.1 (human gut metagenome assembly).

[0294] **SEQ ID NO: 159** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO:158).

[0295] **SEQ ID NO: 160** is an exemplary Cas13d protein sequence from contig emb |0HCP01000044.1 (human gut metagenome assembly).

[0296] **SEQ ID NO: 161** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 160).

[0297] **SEQ ID NO: 162** is an exemplary Cas13d protein sequence from contig embl0GDF01008514.1 | (human gut metagenome assembly).

[0298] **SEQ ID NO: 163** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 162).

[0299] **SEQ ID NO: 164** is an exemplary Cas13d protein sequence from contig emb|OGPN01002610.1 (human gut metagenome assembly).

[0300] **SEQ ID NO: 165** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 164).

[0301] **SEQ ID NO: 166** is an exemplary Cas13d protein sequence from contig NFIR01000008.1 (*Eubacterium* sp. An3, from chicken gut metagenome).

[0302] **SEQ ID NO: 167** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 166).

[0303] **SEQ ID NO: 168** is an exemplary Cas13d protein sequence from contig NFLV01000009.1 (*Eubacterium* sp. An11 from chicken gut metagenome).

[0304] **SEQ ID NO: 169** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 168).

[0305] **SEQ ID NOS: 171-174** are an exemplary Cas13d motif sequences.

[0306] **SEQ ID NO: 175** is an exemplary Cas13d protein sequence from contig OJMM01002900 human gut metagenome sequence.

[0307] **SEQ ID NO: 176** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 175).

[0308] **SEQ ID NO: 177** is an exemplary Cas13d protein sequence from contig ODAI011611274.1 gut metagenome sequence.

[0309] **SEQ ID NO: 178** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 177).

[0310] **SEQ ID NO: 179** is an exemplary Cas13d protein sequence from contig OIZX01000427.1.

[0311] **SEQ ID NO: 180** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 179).

[0312] **SEQ ID NO: 181** is an exemplary Cas13d protein sequence from contig emb|OCVV012889144.1|.

[0313] **SEQ ID NO: 182** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 181).

[0314] **SEQ ID NO: 183** is an exemplary Cas13d protein sequence from contig OCTW011587266.1

[0315] **SEQ ID NO: 184** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 183).

[0316] **SEQ ID NO: 185** is an exemplary Cas13d protein sequence from contig emb |OGNFO 1009141.1.

[0317] **SEQ ID NO: 186** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 185).

[0318] **SEQ ID NO: 187** is an exemplary Cas13d protein sequence from contig emb |OIEN01002196.1.

[0319] **SEQ ID NO: 188** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 187).

[0320] **SEQ ID NO: 189** is an exemplary Cas13d protein sequence from contig e-k87\_11092736.

[0321] **SEQ ID NOS: 190-193** are exemplary consensus DR nucleic acid sequences (goes with SEQ ID NO: 189).

[0322] **SEQ ID NO: 194** is an exemplary Cas13d sequence from Gut\_metagenome\_contig6893000291.

[0323] **SEQ ID NOS: 195-197** are exemplary Cas13d motif sequences.

[0324] **SEQ ID NO: 198** is an exemplary Cas13d protein sequence from Ga0224415\_10007274.

[0325] **SEQ ID NO: 199** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 198).

[0326] **SEQ ID NO: 200** is an exemplary Cas13d protein sequence from EMG\_10003641.

[0327] **SEQ ID NO: 202** is an exemplary Cas13d protein sequence from Ga0129306\_1000735.

[0328] **SEQ ID NO: 201** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 200).

[0329] **SEQ ID NO: 202** is an exemplary Cas13d protein sequence from Ga0129306\_1000735.

[0330] **SEQ ID NO: 203** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 203)

[0331] **SEQ ID NO: 204** is an exemplary Cas13d protein sequence from GaO129317\_1008067.

[0332] **SEQ ID NO: 205** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 204).

[0333] **SEQ ID NO: 206** is an exemplary Cas13d protein sequence from Ga0224415\_10048792.

[0334] **SEQ ID NO: 207** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 206).

[0335] **SEQ ID NO: 208** is an exemplary Cas13d protein sequence from 160582958\_gene49834.

[0336] **SEQ ID NO: 209** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 208).

[0337] **SEQ ID NO: 210** is an exemplary Cas13d protein sequence from 250twins\_35838\_GL0110300.

[0338] **SEQ ID NO: 211** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 210).

[0339] **SEQ ID NO: 212** is an exemplary Cas13d protein sequence from 250twins\_36050\_GLOI58985.

[0340] **SEQ ID NO: 213** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 212). An exemplary direct repeat sequence for Cas13d Seq212 comprises the nucleic acid sequence: TAGCCCTGCAGTAAGGCAGGGTTCTAAGAC (SEQ ID NO: 12279).

[0341] **SEQ ID NO: 214** is an exemplary Cas13d protein sequence from 31009\_GL0034153.

[0342] **SEQ ID NO: 215** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 214).

[0343] **SEQ ID NO: 216** is an exemplary Cas13d protein sequence from

530373\_GL0023589.

[0344] **SEQ ID NO: 217** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 216).

[0345] **SEQ ID NO: 218** is an exemplary Cas13d protein sequence from BMZ-1B\_GL0037771.

[0346] **SEQ ID NO: 219** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 218).

[0347] **SEQ ID NO: 220** is an exemplary Cas13d protein sequence from BMZ-1B\_GL0037915.

[0348] **SEQ ID NO: 221** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 220).

[0349] **SEQ ID NO: 222** is an exemplary Cas13d protein sequence from BMZ-1B\_GL0069617.

[0350] **SEQ ID NO: 223** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 222).

[0351] **SEQ ID NO: 224** is an exemplary Cas13d protein sequence from DLF014\_GL0011914.

[0352] **SEQ ID NO: 225** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 224).

[0353] **SEQ ID NO: 226** is an exemplary Cas13d protein sequence from EYZ-362B\_GL0088915.

[0354] **SEQ ID NO: 227-228** are exemplary consensus DR nucleic acid sequences (goes with SEQ ID NO: 226).

[0355] **SEQ ID NO: 229** is an exemplary Cas13d protein sequence from Ga009936410024192.

[0356] **SEQ ID NO: 230** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 229).

[0357] **SEQ ID NO: 231** is an exemplary Cas13d protein sequence from Ga0187910\_10006931.

[0358] **SEQ ID NO: 232** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 231).

[0359] **SEQ ID NO: 233** is an exemplary Cas13d protein sequence from Ga0187910\_10015336.

[0360] **SEQ ID NO: 234** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 233).

[0361] **SEQ ID NO: 235** is an exemplary Cas13d protein sequence from Ga0187910\_10040531.

[0362] **SEQ ID NO: 236** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 23).

[0363] **SEQ ID NO: 237** is an exemplary Cas13d protein sequence from Ga0187911\_10069260.

[0364] **SEQ ID NO: 238** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 237).

[0365] **SEQ ID NO: 239** is an exemplary Cas13d protein sequence from MH0288\_GL0082219.

[0366] **SEQ ID NO: 240** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 239).

[0367] **SEQ ID NO: 241** is an exemplary Cas13d protein sequence from O2.UC29-0\_GL0096317.

[0368] **SEQ ID NO: 242** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 241).

[0369] **SEQ ID NO: 243** is an exemplary Cas13d protein sequence from PIG-014\_GL0226364.

[0370] **SEQ ID NO: 244** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 243).

[0371] **SEQ ID NO: 245** is an exemplary Cas13d protein sequence from PIG-018\_GL0023397.

[0372] **SEQ ID NO: 246** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 245).

[0373] **SEQ ID NO: 247** is an exemplary Cas13d protein sequence from PIG-025\_GL0099734.

[0374] **SEQ ID NO: 248** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 247).

[0375] **SEQ ID NO: 249** is an exemplary Cas13d protein sequence from PIG-028\_GL0185479.

[0376] **SEQ ID NO: 250** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 249).

[0377] **SEQ ID NO: 251** is an exemplary Cas13d protein sequence from -Ga0224422\_10645759.

[0378] **SEQ ID NO: 252** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 251).

[0379] **SEQ ID NO: 253** is an exemplary Cas13d protein sequence from ODAI chimera.

[0380] **SEQ ID NO: 254** is an exemplary consensus DR nucleic acid sequence (goes with SEQ ID NO: 253).

[0381] **SEQ ID NO: 255** is an HEPN motif.

[0382] **SEQ ID NOs: 256 and 257** are exemplary Cas13d nuclear localization signal amino acid and nucleic acid sequences, respectively.

[0383] **SEQ ID NOs: 258 and 260** are exemplary SV40 large T antigen nuclear localization signal amino acid and nucleic acid sequences, respectively.

[0384] **SEQ ID NO: 259** is a dCas9 target sequence.

[0385] **SEQ ID NO: 261** is an artificial *Eubacterium siraeum* nCasI array targeting ccdB.

[0386] **SEQ ID NO: 262** is a full 36 nt direct repeat.

[0387] **SEQ ID NOs: 263-266** are spacer sequences.

[0388] **SEQ ID NO: 267** is an artificial uncultured *Ruminoccus* sp. nCasI array targeting ccdB.

[0389] **SEQ ID NO: 268** is a full 36 nt direct repeat.

[0390] **SEQ ID NOs: 269-272** are spacer sequences.

[0391] **SEQ ID NO: 273** is a ccdB target RNA sequence.

[0392] **SEQ ID NOs: 274-277** are spacer sequences.

[0393] **SEQ ID NO: 278** is a mutated Cas13d sequence, NLS-Ga\_0531(trunc)-NLS-HA. This mutant has a deletion of the non-conserved N-terminus.

[0394] **SEQ ID NO: 279** is a mutated Cas13d sequence, NES-Ga\_053l(trunc)-NES-HA. This mutant has a deletion of the non-conserved N-terminus.

[0395] **SEQ ID NO: 280** is a full-length Cas13d sequence, NLS-RfxCas13d-NLS-HA.

[0396] **SEQ ID NO: 281** is a mutated Cas13d sequence, NLS-RfxCas13d(del5)-NLS-HA. This mutant has a deletion of amino acids 558-587.

[0397] **SEQ ID NO: 282** is a mutated Cas13d sequence, NLS-RfxCas13d(del5.12)-NLS-HA. This mutant has a deletion of amino acids 558-587 and 953-966.

[0398] **SEQ ID NO: 283** is a mutated Cas13d sequence, NLS-RfxCas13d(del5.13)-NLS-HA. This mutant has a deletion of amino acids 376-392 and 558-587.

[0399] **SEQ ID NO: 284** is a mutated Cas13d sequence, NLS-RfxCas13d(del5.12+5.13)-NLS-HA. This mutant has a deletion of amino acids 376-392, 558-587, and 953-966.

[0400] **SEQ ID NO: 285** is a mutated Cas13d sequence, NLS-RfxCas13d(dell3)-NLS-HA. This mutant has a deletion of amino acids 376-392.

[0401] **SEQ ID NO: 286** is an effector sequence used to edit expression of ADAR2. Amino acids 1 to 969 are dRfxCas13, aa 970 to 991 are an NLS sequence, and amino acids 992 to 1378 are ADAR2DD.

[0402] **SEQ ID NO: 287** is an exemplary HIV NES protein sequence.

[0403] **SEQ ID NOS: 288-291** are exemplary Cas13d motif sequences.

[0404] **SEQ ID NO: 292** is Cas13d ortholog sequence MH\_4866.

[0405] **SEQ ID NO: 293** is an exemplary Cas13d protein sequence from 037\_-  
\_embloIZA01000315.11

[0406] **SEQ ID NO: 294** is an exemplary Cas13d protein sequence from PIG-  
022\_GL002635.1.

[0407] **SEQ ID NO: 295** is an exemplary Cas13d protein sequence from PIG-  
046\_GL0077813.

[0408] **SEQ ID NO: 296** is an exemplary Cas13d protein sequence from  
pig\_chimera.

[0409] **SEQ ID NO: 297** is an exemplary nuclease-inactive or dead Cas13d (dCas13d) protein sequence from *Ruminococcus flavefaciens* XPD3002 (CasRx)

[0410] **SEQ ID NO: 298** is an exemplary Cas13d protein sequence.

[0411] **SEQ ID NO: 299** is an exemplary Cas13d protein sequence from (contig tpg|DJXD01000002.1| ; uncultivated *Ruminococcus* assembly, UBA7013, from sheep gut metagenome)..

[0412] **SEQ ID NO: 300** is an exemplary Cas13d direct repeat nucleotide sequence from Cas13d (contig tpg|DJXD01000002.1| ; uncultivated *Ruminococcus* assembly, UBA7013, from sheep gut metagenome (goes with SEQ ID NO: 299).

[0413] **SEQ ID NO: 301** is an exemplary Cas13d protein contig emb|OBLI01020244.

[0414] Yan et al. (2018) *Mol Cell*. 70(2):327-339 (doi: 10.1016/j.molcel.2018.02.2018) and Konermann et al. (2018) *Cell* 173(3):665-676 (doi: 10.1016/j.cell/2018.02.033) have described Cas13d proteins and both of which are incorporated by reference herein in their entireties. Also see WO Publication Nos. WO2018/183403 (CasM, which is Cas13d) and WO2019/006471 (Cas13d), which are incorporated herein by reference in their entirety.

[0415] **SEQ ID NO: 303** is an exemplary CasM protein from *Eubacterium siraeum*.

[0416] **SEQ ID NO: 304** is an exemplary CasM protein from *Ruminococcus* sp., isolate 2789STDY5834971.

[0417] **SEQ ID NO: 305** is an exemplary CasM protein from *Ruminococcus bicirculans*.

[0418] **SEQ ID NO: 306** is an exemplary CasM protein from *Ruminococcus* sp., isolate 2789STDY5608892.

[0419] **SEQ ID NO: 307** is an exemplary CasM protein from *Ruminococcus* sp. CAG:57.

[0420] **SEQ ID NO: 308** is an exemplary CasM protein from *Ruminococcus flavefaciens* FD-1.

[0421] **SEQ ID NO: 309** is an exemplary CasM protein from *Ruminococcus albus* strain KH2T6.

[0422] **SEQ ID NO: 310** is an exemplary CasM protein from *Ruminococcus flavefaciens* strain XPD3002.

[0423] **SEQ ID NO: 311** is an exemplary CasM protein from *Ruminococcus* sp., isolate 2789STDY5834894.

- [0424] SEQ ID NO: 312 is an exemplary RtcB homolog.
- [0425] SEQ ID NO: 313 is an exemplary WYL from *Eubacterium siraeum* + C-terminal NLS.
- [0426] SEQ ID NO: 314 is an exemplary WYL from *Ruminococcus* sp. isolate 2789STDY5834971 + C-term NLS.
- [0427] SEQ ID NO: 315 is an exemplary WYL from *Ruminococcus bicirculans* + C-term NLS.
- [0428] SEQ ID NO: 316 is an exemplary WYL from *Ruminococcus* sp. isolate 2789STDY5608892 + C-term NLS.
- [0429] SEQ ID NO: 317 is an exemplary WYL from *Ruminococcus* sp. CAG:57 + C-term NLS.
- [0430] SEQ ID NO: 318 is an exemplary WYL from *Ruminococcus flavefaciens* FD-1 + C-term NLS.
- [0431] SEQ ID NO: 319 is an exemplary WYL from *Ruminococcus albus* strain KH2T6 + C-term NLS.
- [0432] SEQ ID NO: 320 is an exemplary WYL from *Ruminococcus flavefaciens* strain XPD3002 + C-term NLS.
- [0433] SEQ ID NO: 321 is an exemplary RtcB from *Eubacterium siraeum* + C-term NLS.
- [0434] SEQ ID NO: 322 is an exemplary direct repeat sequence of *Ruminococcus flavefaciens* XPD3002 Cas13d (CasRx).
- [0435] Exemplary wild type Cas13d proteins of the disclosure may comprise or consist of the amino acid sequence SEQ ID NO: 92 or SEQ ID NO: 298 (Cas13d protein also known as CasRx).
- [0436] An exemplary direct repeat sequence of *Ruminococcus flavefaciens* XPD3002 Cas13d (CasRx) comprises the nucleic acid sequence:  
AACCCCTACCAACTGGTCGGGTTTGAAAC (SEQ ID NO: 302).

#### ***Inactivated Cas13d proteins***

[0437] The disclosure provides Cas13d protein variants comprising deactivating mutations to produce deactivated Cas13d (dCas13d) variants. In some aspects, deactivating mutations reduce or eliminate the nuclease activity of Cas13d. In some aspects, dCas13d variants comprise at least one, at least two, at least three, at least four, at least five, at least six, at least seven, at least eight, at least nine, or at least ten mutations to a Cas13d sequence that reduce or eliminate the nuclease activity of the Cas13d thereby producing a dCas13d variant of the

Cas13d protein. Deactivated Cas13d variants retain the full ability to bind guide RNAs.

Deactivated Cas13d variants also retain RNA-binding activity.

[0438] SEQ ID NO: 397 is an exemplary dCas13d protein, dSeq42. dSeq42 comprises an H841A mutation of the Cas13d protein comprising SEQ ID NO: 42. A nucleic acid sequence encoding SEQ ID NO: 397 is set forth in SEQ ID NO: 461. An exemplary direct repeat sequence for Cas13d Seq42 comprises the nucleic acid sequence:

CACCTCTGCAAACTGCAGGGGTCTAAAAC (SEQ ID NO: 462).

[0439] SEQ ID NO: 396 is an exemplary dCas13d protein, dSeq212. dSeq212 comprises a H919A mutation of the Cas13d protein comprising SEQ ID NO: 212. A nucleic acid sequence encoding SEQ ID NO: 396 is set forth in SEQ ID NO: 395.

[0440] SEQ ID NO: 398 is an exemplary dCas13d protein, dSeq208. dSeq208 comprises a H901A mutation of the Cas13d protein comprising SEQ ID NO: 208.

[0441] SEQ ID NO: 399 is an exemplary dCas13d protein, dSeq179. dSeq179 comprises a H840A mutation of the Cas13d protein comprising SEQ ID NO: 179.

[0442] SEQ ID NO: 400 is an exemplary dCas13d protein, dSeq198. dSeq198 comprises a H841A mutation of the Cas13d protein comprising SEQ ID NO: 198.

[0443] SEQ ID NO: 401 is an exemplary dCas13d protein, dSeq189. dSeq189 comprises a H892A mutation of the Cas13d protein comprising SEQ ID NO: 189.

[0444] SEQ ID NO: 402 is an exemplary dCas13d protein, dCasRX. dCasRX comprises an R238A, H243A, R857A, and H862A mutation of the Cas13d protein comprising SEQ ID NO: 92. A nucleic acid sequence encoding SEQ ID NO: 402 is set forth in SEQ ID NO: 463.

[0445] SEQ ID NO: 12307 is an exemplary dCas13d protein, dSeq212 4aa. A nucleic acid sequence encoding SEQ ID NO: 12307 is set forth in SEQ ID NO: 12278.

#### ***gRNA Target Sequences***

[0446] The compositions of the disclosure bind a target RNA sequence. In some aspects, the target RNA sequence comprises a pre-mRNA sequence. In some aspects, the target pre-mRNA sequence comprises an exonic sequence to be skipped. In one embodiment, the target pre-mRNA comprises a sequence motif corresponding to a spacer sequence of the guide RNA corresponding to the RNA-guided RNA-binding protein. In some embodiments, one or more spacer sequences are used to target one or more target sequences. In some embodiments, multiple spacers are used to target multiple target RNAs. Such target RNAs can be different target sites within the same RNA molecule or can be different target sites within different RNA molecules. Spacer sequences can also target non-coding RNA. In some embodiments, multiple promoters, e.g., Pol III promoters) can be used to drive expression of

multiple spacers in a gRNA for targeting multiple target RNAs. In one embodiment, the binding of the target RNA(s) or target sequence motif(s) induces altered RNA splicing during mRNA processing. In some aspects, the altered RNA splicing involved the targeted exon to be skipped and not incorporated into the mature mRNA.

**[0447]** In some embodiments of the compositions and methods of the disclosure, the sequence motif of the target RNA is a signature of a disease or disorder.

**[0448]** A sequence motif of the disclosure may be isolated or derived from a sequence of foreign or exogenous sequence found in a genomic sequence, and therefore translated into an mRNA molecule of the disclosure or a sequence of foreign or exogenous sequence found in an RNA sequence of the disclosure.

**[0449]** A target pre-mRNA sequence of the disclosure may comprise, consist of, be situated by, or be associated with a mutation in an endogenous sequence that causes a disease or disorder. The mutation may comprise or consist of a sequence substitution, inversion, deletion, insertion, transposition, or any combination thereof.

**[0450]** In some embodiments, the deletion is a single nucleotide. In some embodiments, the deletion can be one or more nucleotides. In some embodiments, the deletion is an exonic or intronic sequence. In some embodiments, the deletion comprises an intronic and exonic sequence.

**[0451]** In some embodiments, the insertion is a single nucleotide. In some embodiments, the insertion can be one or more nucleotides. In some embodiments, the insertion occurs in an exonic or intronic sequence. In some embodiments, the insertion occurs in an intronic and exonic sequence. A target pre-mRNA sequence of the disclosure may comprise or consist of a repeated sequence. In some embodiments, the repeated sequence may be associated with a microsatellite instability (MSI). MSI at one or more loci results from impaired DNA mismatch repair mechanisms of a cell of the disclosure. A hypervariable sequence of DNA may be transcribed into an mRNA of the disclosure comprising a target sequence comprising or consisting of the hypervariable sequence.

**[0452]** A target pre-mRNA sequence of the disclosure may comprise or consist of a biomarker. The biomarker may indicate a risk of developing a disease or disorder. The biomarker may indicate a healthy gene (low or no determinable risk of developing a disease or disorder). The biomarker may indicate an edited gene. Exemplary biomarkers include, but are not limited to, single nucleotide polymorphisms (SNPs), sequence variations or mutations, epigenetic marks, splice acceptor sites, exogenous sequences, heterologous sequences, and any combination thereof.

[0453] A target pre-mRNA sequence of the disclosure may comprise or consist of a secondary, tertiary or quaternary structure. The secondary, tertiary or quaternary structure may be endogenous or naturally occurring. The secondary, tertiary or quaternary structure may be induced or non-naturally occurring. The secondary, tertiary or quaternary structure may be encoded by an endogenous, exogenous, or heterologous sequence.

[0454] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule comprises or consists of between 2 and 1,000 nucleotides or nucleic acid bases, inclusive of the endpoints. In some embodiments, the target sequence of an RNA molecule comprises or consists of between 2 and 100 nucleotides or nucleic acid bases, inclusive of the endpoints. In some embodiments, the target sequence of an RNA molecule comprises or consists of between 2 and 50 nucleotides or nucleic acid bases, inclusive of the endpoints. In some embodiments, the target sequence of an RNA molecule comprises or consists of between 2 and 20 nucleotides or nucleic acid bases, inclusive of the endpoints. In some embodiments, the target sequence of an RNA molecule comprises or consists of between 20-30 nucleotides or nucleic acid bases, inclusive of the endpoints. In some embodiments, the target sequence of an RNA molecule comprises or consists of about 26 nucleotides or nucleic acid bases, inclusive of the endpoints.

[0455] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule is continuous. In some embodiments, the target sequence of an RNA molecule is discontinuous. For example, the target sequence of an RNA molecule may comprise or consist of one or more nucleotides or nucleic acid bases that are not contiguous because one or more intermittent nucleotides are positioned in between the nucleotides of the target sequence.

[0456] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule is naturally occurring. In some embodiments, the target sequence of an RNA molecule is non-naturally occurring. Exemplary non-naturally occurring target sequences may comprise or consist of sequence variations or mutations, chimeric sequences, exogenous sequences, heterologous sequences, chimeric sequences, recombinant sequences, sequences comprising a modified or synthetic nucleotide or any combination thereof.

[0457] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule binds to a guide RNA of the disclosure. In some embodiments of the compositions and methods of the disclosure, one or more target sequences of an RNA molecule binds to one or more guide RNA spacer sequences of the disclosure.

[0458] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule binds to a first RNA binding protein of the disclosure.

[0459] In some embodiments of the compositions and methods of the disclosure, a target sequence of an RNA molecule binds to a second RNA binding protein of the disclosure.

[0460] Compositions of the disclosure comprise a gRNA comprising a spacer sequence that specifically binds to a pre-mRNA sequence comprising a mutation resulting in a disease or disorder. In some embodiments, the mutation occurs in an exon of the pre-mRNA. In some embodiments, the mutation occurs in an intron of the pre-mRNA. In some aspects, the pre-mRNA sequence comprises at least one of: an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), and a polypyrimidine tract (PPY). In some aspects, the 5' splice site is a splice donor site. In some aspects, the 3' splice site is a splice acceptor site. In some aspects, the pre-mRNA sequence comprises an exon splicing silencer (ESS) motif or an intron splicing silencer (ISS) motif.

[0461] In some embodiments, the spacer which binds the target pre-mRNA sequence comprises or consists of about 20-30 nucleotides. In some embodiments, a gRNA comprises one or more spacer sequences.

[0462] Exemplary gRNA spacer sequences of the disclosure can be configured to specifically bind to an RNA sequence encoding any gene in a subject's genome. Exemplary gRNA spacer sequences of the disclosure can be configured to specifically bind to a pre-mRNA sequence encoding any gene, or derived from any gene, in a subject's genome.

[0463] Exemplary gRNA spacer sequences of the disclosure specifically bind to a USH2A pre-mRNA. In some embodiments, the sequence encoding for USH2A comprises a mutation. In some embodiments, the USH2A pre-mRNA is human. In some embodiments, the USH2A pre-mRNA is murine.

[0464] In some embodiments, the gRNA spacer sequence targets an USH2A sequence comprising a c.2299delG mutation. In some embodiments, the nucleic acid sequence encoding wild-type human USH2A mRNA comprises or consists of SEQ ID NO: 465 or SEQ ID NO: 466. In some embodiments, the nucleic acid sequence encoding wild-type human USH2A pre-mRNA comprises or consists of SEQ ID NO: 467. In some embodiments, the amino acid sequence encoding human usherin comprises or consists of SEQ ID NO: 468. In some embodiments, the amino acid sequence encoding human usherin comprises a C759F mutation.

[0465] In some embodiments, the nucleic acid sequence encoding wild-type murine USH2A mRNA comprises or consists of SEQ ID NO: 486. In some embodiments, the amino acid sequence encoding murine usherin comprises or consists of SEQ ID NO: 488.

[0466] Exemplary gRNA spacer sequences of the disclosure specifically bind to a DMD pre-mRNA. In some embodiments, the sequence encoding for DMD comprises a mutation. In some embodiments, the DMD pre-mRNA is human.

[0467] In some embodiments, the nucleic acid sequence encoding wild-type human DMD mRNA comprises or consists of SEQ ID NO: 470. In some embodiments, the nucleic acid sequence encoding wild-type human DMD pre-mRNA comprises or consists of SEQ ID NO: 471. In some embodiments, the nucleic acid sequence encoding human DMD comprises a deletion. In some embodiments, the deletion is a single nucleotide. In some embodiments, the deletion can be one or more nucleotides. In some embodiments, the deletion is an exonic or intronic sequence. In some embodiments, the deletion comprises an intronic and exonic sequence. In some embodiments, the deletion occurs across a region comprising one or more of exon 45, exon 46, exon 47, exon 48, exon 49, exon 50, exon 51, exon 52, exon 53, exon 54, and exon 55 and any introns upstream downstream of therebetween of these exons. In some embodiments, the amino acid sequence encoding human dystrophin comprises or consists of SEQ ID NO: 472.

[0468] Exemplary gRNA spacer sequences of the disclosure specifically bind to a MAPT pre-mRNA. In some embodiments, the sequence encoding for MAPT comprises a mutation. In some embodiments, the MAPT pre-mRNA is human.

[0469] In some embodiments, the nucleic acid sequence encoding wild-type human MAPT mRNA comprises or consists of SEQ ID NO: 478. In some embodiments, the nucleic acid sequence encoding MAPT comprises a mutation in exon 10 or intron 10. In some embodiments, the nucleic acid sequence encoding wild-type human MAPT pre-mRNA comprises or consists of SEQ ID NO: 479. In some embodiments, the amino acid sequence encoding human tau comprises or consists of SEQ ID NO: 480.

[0470] Exemplary gRNA spacer sequences of the disclosure specifically bind to a SOD1 pre-mRNA. In some embodiments, the sequence encoding for SOD1 comprises a mutation. In some embodiments, the SOD1 pre-mRNA is human.

[0471] In some embodiments, the nucleic acid sequence encoding wild-type human SOD1 mRNA comprises or consists of SEQ ID NO: 482. In some embodiments, the nucleic acid sequence encoding wild-type human SOD1 pre-mRNA comprises or consists of SEQ ID NO:

483. In some embodiments, the amino acid sequence encoding human superoxide dismutase 1 (SOD1) comprises or consists of SEQ ID NO: 484.

[0472] Exemplary gRNA spacer sequences of the disclosure specifically bind to a LRRK2 pre-mRNA. In some embodiments, the sequence encoding for LRRK2 comprises a mutation. In some embodiments, the LRRK2 pre-mRNA is human.

[0473] In some embodiments, the nucleic acid sequence encoding wild-type human LRRK2 mRNA comprises or consists of SEQ ID NO: 474. In some embodiments, the nucleic acid sequence encoding wild-type human LRRK2 pre-mRNA comprises or consists of SEQ ID NO: 475. In some embodiments, the amino acid sequence encoding Leucine-rich repeat kinase 2 (LRRK2) comprises or consists of SEQ ID NO: 476. In some embodiments, the amino acid sequence encoding human LRRK2 comprises a mutation comprising one or more of G2019S, R1441G, R1441C, R1441H, Y1699C, and I2020T.

[0474] Exemplary gRNA spacer sequences of the disclosure specifically bind to a SMN1 pre-mRNA. In some embodiments, the sequence encoding for SMN1 comprises a mutation. In some embodiments, the SMN1 pre-mRNA is human.

[0475] In some embodiments, the nucleic acid sequence encoding wild-type human SMN1 mRNA comprises or consists of SEQ ID NO: 12324. In some embodiments, the nucleic acid sequence encoding wild-type human SMN1 gene is set forth in ENST00000380707.9. In some embodiments, the amino acid sequence encoding human survival motor neuron 1 (SMN1) comprises or consists of SEQ ID NO: 12322.

[0476] Exemplary gRNA spacer sequences of the disclosure specifically bind to a SMN2 pre-mRNA. In some embodiments, the sequence encoding for SMN2 comprises a mutation. In some embodiments, the SMN2 pre-mRNA is human.

[0477] In some embodiments, the nucleic acid sequence encoding wild-type human SMN2 mRNA comprises or consists of SEQ ID NO: 12325. In some embodiments, the nucleic acid sequence encoding wild-type human SMN2 gene is set forth in ENST00000380743.9. In some embodiments, the amino acid sequence encoding human survival motor neuron 2 (SMN2) comprises or consists of SEQ ID NO: 12323.

#### **AAV vectors**

[0478] In some embodiments of the compositions of the disclosure, pre-mRNA-targeting compositions are packaged as AAV vectors. In some embodiments of the compositions of the disclosure, pre-mRNA-targeting Cas13d compositions are packaged as AAV vectors. In some embodiments of the compositions of the disclosure, pre-mRNA-targeting PUF/PUMBY compositions are packaged as AAV vectors.

[0479] In some embodiments, RNA-targeting Cas13d vectors comprise a nucleic acid sequence encoding at least one guide RNA under control of a promoter and a sequence encoding a cas13d peptide under the control of a second promoter.

[0480] In some embodiments, an exemplary pre-mRNA-targeting Cas13d composition comprises from 5' to 3': a) a promoter, b) a cas13d gRNA, wherein the gRNA comprises i) a direct repeat sequence and ii) a RNA targeting spacer sequence, c) a promoter, and d) a sequence encoding dCas13d. In some embodiments, the composition further comprises a 5' ITR sequence and a 3' ITR sequence.

[0481] In some embodiments, an exemplary pre-mRNA-targeting Cas13d composition comprises from 5' to 3': a) a promoter, b) a first cas13d gRNA, wherein the gRNA comprises i) a direct repeat sequence and ii) a RNA targeting spacer sequence, c) a second cas13d gRNA, wherein the gRNA comprises i) a direct repeat sequence and ii) a RNA targeting spacer sequence, d) a promoter, and e) a sequence encoding dCas13d. In some embodiments, the composition further comprises a 5' ITR sequence and a 3' ITR sequence.

[0482] In some embodiments, an exemplary pre-mRNA-targeting Cas13d composition comprises from 5' to 3': a) a first promoter, b) a first cas13d gRNA, wherein the gRNA comprises i) a direct repeat sequence and ii) a RNA targeting spacer sequence, c) a second promoter, d) a second cas13d gRNA, wherein the gRNA comprises i) a direct repeat sequence and ii) a RNA targeting spacer sequence, e) a promoter, and f) a sequence encoding dCas13d. In some embodiments, the composition further comprises a 5' ITR sequence and a 3' ITR sequence.

[0483] In some embodiments, an exemplary pre-mRNA-targeting PUF composition comprises from 5' to 3': a) a promoter, b) a sequence encoding a PUF protein. In some embodiments, the composition further comprises a 5' ITR sequence and a 3' ITR sequence.

[0484] In some embodiments, an exemplary pre-mRNA-targeting Cas13d composition includes additional elements such as signal sequences, regulatory sequences, linkers, tags, and poly A sequences. For instance, in one embodiment, a pre-mRNA-targeting Cas13d composition comprises from 5' to 3': a) a human U6 promoter, b) a cas13d gRNA, wherein the gRNA comprises i) a direct repeat sequence and ii) a RNA targeting spacer sequence, c) a promoter, d) a kozak sequence, an e) NLS sequence, f) a linker sequence, g) a sequence encoding dCas13d, h) a linker sequence, i) an NLS sequence, j) a linker sequence, k) a tag sequence, and l) a poly A sequence. In some embodiments, the composition further comprises a 5' ITR sequence and a 3' ITR sequence.

**[0485]** An "AAV vector" as used herein refers to a vector comprising, consisting essentially of, or consisting of one or more nucleic acid molecules and one or more AAV inverted terminal repeat sequences (ITRs). In some aspects, the nucleic acid molecule encodes for a repeat targeting protein and/or composition of the disclosure. Such AAV vectors can be replicated and packaged into infectious viral particles when present in a host cell that provides the functionality of rep and cap gene products, for example, by transfection of the host cell. In some aspects, AAV vectors contain a promoter, at least one nucleic acid that may encode at least one protein or RNA, and/or an enhancer and/or a terminator within the flanking ITRs that is packaged into the infectious AAV particle. The encapsidated nucleic acid portion may be referred to as the AAV vector genome. Plasmids containing AAV vectors may also contain elements for manufacturing purposes, e.g., antibiotic resistance genes, origin of replication sequences etc., but these are not encapsidated and thus do not form part of the AAV particle.

**[0486]** In some aspects, an AAV vector can comprise at least one nucleic acid molecule encoding a repeat targeting composition of the disclosure. In some aspects, an AAV vector can comprise at least one regulatory sequence. In some aspects, an AAV vector can comprise at least one AAV inverted terminal (ITR) sequence. In some aspects, an AAV vector can comprise a first ITR sequence and a second ITR sequence. In some aspects, an AAV vector can comprise at least one promoter sequence. In some aspects, an AAV vector can comprise at least one enhancer sequence. In some aspects, an AAV vector can comprise at least one polyA sequence. In some aspects, an AAV vector can comprise at least one linker sequence. In some aspects, an AAV vector of the disclosure can comprise at least one nuclear localization signal, or nuclear export signal and or both. In some aspects, an AAV vector of the disclosure will contain a WPRE (woodchuck hepatitis virus post-transcriptional regulatory element). In some aspects, an AAV vector of the disclosure can comprise a Cas protein, peptide, or fragment thereof. In some aspects, an AAV vector of the disclosure can comprise an endonuclease protein, peptide, or fragment thereof. In some aspects, an AAV vector of the disclosure can comprise a guide RNA, in some cases a pre-mRNA targeting guide RNA. In some aspects, AAV vectors of the disclosure can comprise a fusion protein comprising one or more elements of the disclosure, including, but not limited to, an RNA-targeting protein (such as a Cas, PUF, or PUMBY) and an endonuclease. Optionally, fusion proteins of the AAV vector can further comprise a linker amino acid sequence between the one or more elements of the disclosure.

[0487] In some aspects, an AAV vector can comprise a first AAV ITR sequence, a promoter sequence, a RNA-targeting composition nucleic acid molecule, a regulatory sequence and a second AAV ITR sequence. In some aspects, an AAV vector can comprise, in the 5' to 3' direction, a first AAV ITR sequence, a promoter sequence, a transgene nucleic acid molecule, and a second AAV ITR sequence.

[0488] In some embodiments of the compositions and methods of the disclosure, a vector comprises a guide RNA of the disclosure. In some embodiments, the vector comprises at least one guide RNA of the disclosure. In some embodiments, the vector comprises one or more guide RNA(s) of the disclosure. In some embodiments, the vector comprises two or more guide RNAs of the disclosure. In one embodiment, the vector comprises three guide RNAs. In one embodiment, the vector comprises four guide RNAs. In some embodiments, the vector comprises a PolIII promoter, one or multiple guides, a PolII promoter, the Cas protein, a regulatory element and a polyA. In some embodiments, the vector further comprises a guided or non-guided RNA-binding protein of the disclosure. In some embodiments, the vector further comprises an RNA-binding fusion protein of the disclosure. In some embodiments, the fusion protein comprises a first RNA binding protein and a second RNA binding protein. In some embodiments, the RNA-guided RNA-binding systems comprising an RNA-binding protein and a gRNA are in a single vector. In a particular embodiment, the single vector comprises the RNA-guided RNA-binding systems which are Cas13d RNA-guided RNA-binding systems or catalytic deactivated Cas13d (dCas13d) RNA-guided RNA-binding systems. In one embodiment, the single vector comprises the Cas13d RNA-guided RNA-binding systems which are CasRx or dCasRx RNA-guided RNA-binding systems. In another embodiment, the single vector comprises a non-guided RNA-binding system comprising a PUF or PUMBY-based protein. In another embodiment, the single vector comprises a non-guided RNA-binding system comprising a PUF or PUMBY-based protein fused with a nuclease domain from ZC3H12A, such as E17. In another embodiment, the single vector comprises a dCas13d RNA-binding system fused with a nuclease domain from ZC3H12A, such as E17 (SEQ ID NO: 359).

[0489] In some aspects, the nuclease domain from ZC3H12A, E17 can comprise, consist essentially of, or consist of the amino acid sequence:

GGGTPKAPNLEPPLPEEEKEGSDLRPVVIDGSNVAMSHGNKEVFSCRGILLAVNWFL  
 ERGHTDITVFVPSWRKEQRPDPVITDQHILRELEKKKILVFTPSRRVGGKRVVCYDD  
 RFIVKLAYESDGIVVSNDTYRDLQGERQEWKRFIEERLLMYSFVNDKFMPPDDPLGR  
 HGPSLDNFLRKKPLTLE (SEQ ID NO: 12308).

[0490] In some aspects, the nuclease domain from ZC3H12A, E17 can comprise, consist essentially of, or consist of the amino acid sequence:

SGPCGEKPVLEASPTMSLWFEEDSHSRQGTTPRGQELAAEEASALELQMKVDFFRKL  
 GYSSTEIHSVLQKLGVQADTNTVLGELVKHGTATERERQTSPDPCPQLPLVPRGGGT  
 PKAPNLEPPLPEEEKEGSDLRPVVIDGSNVAMSHGNKEVFSCRGILLAVNWFLEIGH  
 TDITVFVPSWRKEQPRPDVPITDQHILRELEKKKILVFTPSRRVGGKRVVCYDDRFIV  
 KLAYESDGIVVSNPTYRDLQGERQEWKRFIEERLLMYSFVNDKFMPPDDPLGRHGP  
 SLDNFLRKKPLTLEHRKQPCPYGRKCTYGIKCRFFHPERPSCPQRSVADELARANALLS  
 PPRAPSKDKNGRRPSPSSQSSSLLTESEQCSLDGKKLGAQASPGSRQEGLTQTYAPSG  
 RSLAPSGGSGSSFGPTDWLPQTLDLPYVSQDCLDSGIGSLESQMSELWGVRRGGGPG  
 EPGPPRAPHYTGYSPTYGSELPATAAFSAFGRAMGAGHFSVPADYPPAPPAFPFPREYWS  
 EPYPLPPPTSVLQEPPVQSPGAGRSPWGRAGSLAKEQASVYTKLCGVFPFPHLVEAVM  
 GRFPQLLDPQQLAAEILSYKSQHPSE (SEQ ID NO: 12309).

[0491] In some embodiments of the compositions and methods of the disclosure, a first vector comprises a guide RNA of the disclosure and a second vector comprises an RNA-binding protein or RNA-binding fusion protein of the disclosure. In some embodiments, the first vector comprises at least one guide RNA of the disclosure. In some embodiments, the first vector comprises one or more guide RNA(s) of the disclosure. In some embodiments, the first vector comprises two or more guide RNA(s) of the disclosure. In some embodiments, the fusion protein comprises a first RNA binding protein and a second RNA binding protein. In some embodiments, the first vector and the second vector are identical vectors or vector serotypes. In some embodiments, the first vector and the second vector are not identical vectors or vector serotypes. In some embodiments of the compositions and methods of the disclosure, the RNA-binding systems capable of targeting pre-mRNA sequences are in a single vector.

[0492] One type of vector is a "plasmid," which refers to a circular double stranded DNA loop into which additional DNA segments can be inserted, such as by standard molecular cloning techniques. Another type of vector is a viral vector, wherein virally -derived DNA or RNA sequences are present in the vector for packaging into a virus (e.g., retroviruses, replication defective retroviruses, adenoviruses, replication defective adenoviruses, and adeno-associated viruses). Viral vectors also include polynucleotides carried by a virus for transfection into a host cell. In some embodiments, the vector is a lentivirus (such as an integration-deficient lentiviral vector) or adeno-associated viral (AAV) vector. Vectors are capable of autonomous replication in a host cell into which they are introduced such as e.g.,

bacterial vectors having a bacterial origin of replication and episomal mammalian vectors and other vectors such as, e.g., non-episomal mammalian vectors, are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome.

**[0493]** In some embodiments, vectors such as e.g., expression vectors, are capable of directing the expression of genes to which they are operatively-linked. Common expression vectors are often in the form of plasmids. In some embodiments, recombinant expression vectors comprise a nucleic acid provided herein such as e.g., a guide RNA which can be expressed from a DNA sequence, and a nucleic acid encoding a Cas 13d protein, in a form suitable for expression of a protein in a host cell. Recombinant expression vectors include one or more regulatory elements, which may be selected on the basis of the host cells to be used for expression, that is operatively-linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory element(s) in a manner that allows for expression of the nucleotide sequence such as e.g., in an in vitro transcription/translation system or in a host cell when the vector is introduced into the host cell. Certain embodiments of a vector depend on factors such as the choice of the host cell to be transformed, and the level of expression desired. A vector can be introduced into host cells to thereby produce transcripts, proteins, or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein such as, e.g., CRISPR transcripts, proteins, enzymes, mutant forms thereof, fusion proteins thereof, etc.

#### *AAV Viral Vectors*

**[0494]** In some embodiments of the compositions and methods of the disclosure, a vector of the disclosure is a viral vector. In some embodiments, the viral vector comprises a sequence isolated or derived from a retrovirus. In some embodiments, the viral vector comprises a sequence isolated or derived from a lentivirus. In some embodiments, the viral vector comprises a sequence isolated or derived from an adenovirus. In some embodiments, the viral vector comprises a sequence isolated or derived from an adeno-associated virus (AAV). In some embodiments, the viral vector is replication incompetent. In some embodiments, the viral vector is isolated or recombinant. In some embodiments, the viral vector is self-complementary.

**[0495]** The term "adeno-associated virus" or "AAV" as used herein refers to a member of the class of viruses associated with this name and belonging to the genus Dependoparvovirus, family Parvoviridae. Adeno-associated virus is a single-stranded DNA virus that grows in

cells in which certain functions are provided by a co-infecting helper virus. General information and reviews of AAV can be found in, for example, Carter, 1989, Handbook of Parvoviruses, Vol. 1, pp. 169- 228, and Berns, 1990, Virology, pp. 1743-1764, Raven Press, (New York). It is fully expected that the same principles described in these reviews will be applicable to additional AAV serotypes characterized after the publication dates of the reviews because it is well known that the various serotypes are quite closely related, both structurally and functionally, even at the genetic level. (See, for example, Blacklowe, 1988, pp. 165-174 of Parvoviruses and Human Disease, J. R. Pattison, ed.; and Rose, Comprehensive Virology 3: 1-61 (1974)). For example, all AAV serotypes apparently exhibit very similar replication properties mediated by homologous rep genes; and all bear three related capsid proteins such as those expressed in AAV2. The degree of relatedness is further suggested by heteroduplex analysis which reveals extensive cross-hybridization between serotypes along the length of the genome; and the presence of analogous self-annealing segments at the termini that correspond to "inverted terminal repeat sequences" (ITRs). The similar infectivity patterns also suggest that the replication functions in each serotype are under similar regulatory control. Multiple serotypes of this virus are known to be suitable for gene delivery; all known serotypes can infect cells from various tissue types.

[0496] AAV possesses unique features that make it attractive as a vector for delivering foreign DNA to cells, for example, in gene therapy. AAV infection of cells in culture is noncytopathic, and natural infection of humans and other animals is silent and asymptomatic. Moreover, AAV infects many mammalian cells allowing the possibility of targeting many different tissues in vivo. Moreover, AAV transduces slowly dividing and non-dividing cells, and can persist essentially for the lifetime of those cells as a transcriptionally active nuclear episome (extrachromosomal element). The AAV proviral genome is inserted as cloned DNA in plasmids, which makes construction of recombinant genomes feasible. Furthermore, because the signals directing AAV replication and genome encapsidation are contained within the ITRs of the AAV genome, some or all of the internal approximately 4.3 kb of the genome (encoding replication and structural capsid proteins, rep-cap) may be replaced with foreign DNA to generate AAV vectors. The rep and cap proteins may be provided in trans. Another significant feature of AAV is that it is an extremely stable and hearty virus. It easily withstands the conditions used to inactivate adenovirus (56° to 65°C for several hours), making cold preservation of AAV less critical. AAV may even be lyophilized. Finally, AAV-infected cells are not resistant to superinfection.

[0497] Recombinant AAV (rAAV) genomes of the invention comprise, consist essentially of, or consist of a nucleic acid molecule encoding a pre-mRNA targeting composition (such as a guided RNA-binding compositions such as a Cas13d polypeptide or a non-guided RNA-binding composition such as a PUF or PUMBY polypeptide) and one or more AAV ITRs flanking the nucleic acid molecule. Production of pseudotyped rAAV is disclosed in, for example, WO2001083692. Other types of rAAV variants, for example rAAV with capsid mutations, are also contemplated. *See, e.g.,* Marsic et al., *Molecular Therapy*, 22(11): 1900-1909 (2014). The nucleotide sequences of the genomes of various AAV serotypes are known in the art.

[0498] In some embodiments of the compositions and methods of the disclosure, the viral vector comprises a sequence isolated or derived from an adeno-associated virus (AAV). In some embodiments, the viral vector comprises an inverted terminal repeat sequence or a capsid sequence that is isolated or derived from an AAV of serotype AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10 (AAVrh10), AAV11 or AAV12. In some embodiments, the AAV serotype is AAVrh.74. In one embodiment, the AAV vector comprises a modified capsid. In one embodiment the AAV vector is an AAV2-Tyr mutant vector. In one embodiment the AAV vector comprises a capsid with a non-tyrosine amino acid at a position that corresponds to a surface-exposed tyrosine residue in position Tyr252, Tyr272, Tyr275, Tyr281, Tyr508, Tyr612, Tyr704, Tyr720, Tyr730 or Tyr673 of wild-type AAV2. *See also* WO 2008/124724 incorporated herein in its entirety. In some embodiments, the AAV vector comprises an engineered capsid. AAV vectors comprising engineered capsids include without limitation, AAV2.7m8, AAV9.7m8, AAV2 2tYF, and AAV8 Y733F). In some embodiments, the capsid is a ubiquitination resistant capsid. In another embodiment, the ubiquitination capsid is an AAV2 capsid comprising tyrosine (Y) and serine (S) mutations. In another embodiment, the AAV2 capsid comprises Y, S and threonine (T) mutations. In another embodiment, the AAV2 capsid includes, without limitation, AAV2 capsid mutants such as T455V, T491V, T550V, T659V, Y444+500+730F, and Y444+500+730F+T491V. In some embodiments, the viral vector is replication incompetent. In some embodiments, the viral vector is isolated or recombinant (rAAV). In some embodiments, the viral vector is self-complementary (scAAV).

#### ***AAV ITR SEQUENCES***

[0499] In some embodiments of the compositions and methods of the disclosure, an AAV inverted terminal repeat sequence can comprise any AAV ITR sequence known in the art. In some aspects, an AAV ITR sequence can comprise or consist of an AAV1 ITR sequence, an

AAV2 ITR sequence, an AAV3 ITR sequence, an AAV4 ITR sequence, an AAV5 ITR sequence, an AAV6 ITR sequence, an AAV7 ITR sequence, an AAV8 ITR sequence, an AAV9 ITR sequence, an AAV10 ITR sequence, an AAVrh10 ITR sequence, an AAV11 ITR sequence, an AAV12 ITR sequence, an AAV13 ITR sequence, or an AAVrh74 ITR sequence.

**[0500]** In some aspects, the ITR sequence can comprise a modified AAV ITR sequence.

**[0501]** In some aspects, an AAV ITR sequence can comprise, consist essentially of, or consist of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 458 or 459.

**[0502]** In some aspects, a modified ITR sequence can comprise consist essentially of, or consist of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 458 or 459.

**[0503]** In some aspects, a first AAV ITR sequence can comprise, consist essentially of, or consist of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 458 and a second AAV ITR sequence can comprise, consist essentially of, or consist of a nucleic acid sequence at least 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% (or any percentage in between) identical to SEQ ID NO: 459. In some aspects, the first AAV ITR sequence is positioned at the 5' of a AAV vector. In some aspects, the second AAV ITR sequence is positioned at the 3' of a AAV vector.

**[0504]** In some embodiments of the compositions and methods of the disclosure, a vector of the disclosure is a non-viral vector. In some embodiments, the vector comprises or consists of a nanoparticle, a micelle, a liposome or lipoplex, a polymersome, a polyplex or a dendrimer. In some embodiments, the vector is an expression vector or recombinant expression system. As used herein, the term “recombinant expression system” refers to a genetic construct for the expression of certain genetic material formed by recombination.

#### ***Promoter Sequences***

**[0505]** In some embodiments of the compositions and methods of the disclosure, an expression vector, viral vector or non-viral vector provided herein, includes without limitation, an expression control element. An “expression control element” as used herein refers to any sequence that regulates the expression of a coding sequence, such as a gene. Exemplary expression control elements include but are not limited to promoters, enhancers, microRNAs, post-transcriptional regulatory elements, polyadenylation signal sequences, and introns. Expression control elements may be constitutive, inducible, repressible, or tissue-

specific, for example. A “promoter” is a control sequence that is a region of a polynucleotide sequence at which initiation and rate of transcription are controlled. It may contain genetic elements at which regulatory proteins and molecules may bind such as RNA polymerase and other transcription factors. In some embodiments, expression control by a promoter is tissue-specific. In some embodiments, expression control by a promoter is constitutive or ubiquitous. Non-limiting exemplary promoters include a Pol III promoter such as, e.g., U6 and H1 promoters and/or a Pol II promoter e.g., SV40, CMV (optionally including the CMV enhancer), RSV (Rous Sarcoma Virus LTR promoter (optionally including RSV enhancer), CBA (hybrid CMV enhancer/ chicken  $\beta$ -actin), CAG (hybrid CMV enhancer fused to chicken  $\beta$ -actin), truncated CAG, Cbh (hybrid CBA), EF-1a (human elongation factor alpha-1) or EFS (short intron-less EF-1 alpha), PGK (phosphoglycerol kinase), CEF (chicken embryo fibroblasts), UBC (ubiquitinC), GUSB (lysosomal enzyme beta-glucuronidase), UCOE (ubiquitous chromatin opening element), hAAT (alpha-1 antitrypsin), TBG (thyroxine binding globulin), Desmin (full-length (SEQ ID NO: 655) or truncated (SEQ ID NO: 656)), MCK (muscle creatine kinase), C5-12 (synthetic muscle promoter), CK8e (creatin kinase 8), NSE (neuron-specific enolase), Synapsin, Synapsin-1 (SYN-1), opsin, PDGF (platelet-derived growth factor), PDGF-A, MecP2 (methyl CpG-binding protein 2), CaMKII (Calcium/ Calmodulin-dependent protein kinase II), mGluR2 (metabotropic glutamate receptor 2), NFL (neurofilament light), NFH (neurofilament heavy),  $\eta\beta 2$ , PPE (rat preproenkephalin), ENK (preproenkephalin), Preproenkephalin-neurofilament chimeric promoter, EAAT2 (glutamate transporter), GFAP (glial fibrillary acidic protein), MBP (myelin basic protein), human rhodopsin kinase promoter (hGRK1),  $\beta$ -actin promoter, dihydrofolate reductase promoter, MHCK7 (hybrid promoter of enhancer/ promoter regions of muscle creatine kinase and alpha myosin heavy-chain genes) and combinations thereof. An “enhancer” is a region of DNA that can be bound by activating proteins to increase the likelihood or frequency of transcription. Non-limiting exemplary enhancers and posttranscriptional regulatory elements include the CMV enhancer, MCK enhancer, R-U5' segment in LTR of HTLV-1, SV40 enhancer, the intron sequence between exons 2 and 3 of rabbit  $\beta$ -globin, and Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPPE). In some embodiments, a nucleic acid sequence encoding a WPPE element is set forth in SEQ ID NO: 456.

**[0506]** In some embodiments an intron is used to enhance promoter activity such as a UBB intron. In some embodiments, the UBB intron is used with an EFS promoter.

**[0507]** Tissue-specific ocular promoters include, without limitation, human blue opsin (e.g., HB570 and HB569), human red cone opsin (e.g., PR2.1, PR1.7, PRO.5, 3LCR-PRO.5), human green red opsin (G1.7p), human rhodopsin kinase (RHOK/GRK), human interphotoreceptor retinoid binding protein (hIRBP), human retinal pigment epithelial (hRPE65p, NA65p, SynPIII), bestrophin1 promoter (VMD2), human gamma-synuclein promoter (hSNCGp), synapsin-1 (Syn1), human neurofilament heavy gene promoter (Nefh), human cone arrestin (e.g., CAR/ARR3), human retinal X-arrestin, homeobox protein (CHx10), and glial fibrillary acidic protein (GFAP).

**[0508]** In some embodiments, tissue-specific ocular promoters are capable of driving expression to a retinal cell such as, without limitation, a photoreceptor cell, a retinal pigment epithelial cell, a ganglion cell, an amacrine cell, a bipolar cell, a horizontal cell, a Müller glial cell, a rod cell, or a cone cell.

**[0509]** In some embodiments of the compositions and methods of the disclosure, an expression vector, viral vector or non-viral vector provided herein, includes without limitation, vector elements such as an IRES or 2A peptide sites for configuration of “multicistronic” or “polycistronic” or “bicistronic” or tricistronic” constructs, i.e., having double or triple or multiple coding areas or exons, and as such will have the capability to express from mRNA two or more proteins from a single construct. Multicistronic vectors simultaneously express two or more separate proteins from the same mRNA. The two strategies most widely used for constructing multicistronic configurations are through the use of an IRES or a 2A self-cleaving site. An “IRES” refers to an internal ribosome entry site or portion thereof of viral, prokaryotic, or eukaryotic origin which are used within polycistronic vector constructs. In some embodiments, an IRES is an RNA element that allows for translation initiation in a cap-independent manner. The term “self-cleaving peptides” or “sequences encoding self-cleaving peptides” or “2A self-cleaving site” refer to linking sequences which are used within vector constructs to incorporate sites to promote ribosomal skipping and thus to generate two polypeptides from a single promoter, such self-cleaving peptides include without limitation, T2A, and P2A peptides or other sequences encoding the self-cleaving peptides.

**[0510]** Exemplary vector configurations comprise a promoter or regulatory sequence (promoter/enhancer combination) driving the expression of the nucleic acid encoding the RNA-targeting Cas13d-based system. In another embodiment, a vector configuration comprises a promoter driving expression of the RNA-guided Cas RNase RNA-binding protein, or dCas protein fusion in operable linkage with a second promoter driving expressing

of a cognate gRNA. In another embodiment, the vector configuration comprises a linker and one or more tags.

**[0511]** In some embodiments of the compositions of the disclosure, vectors comprising guide RNA sequences of the disclosure comprises a promoter sequence to drive expression of the guide RNA. In some embodiments, a vector comprising a guide RNA sequence of the disclosure comprises a promoter sequence to drive expression of the guide RNA. In some embodiments, the promoter to drive expression of the guide RNA is a constitutive promoter. In some embodiments, the promoter sequence is an inducible promoter. In some embodiments, the promoter is a sequence is a tissue-specific and/or cell-type specific promoter. In some embodiments, the promoter is a hybrid or a recombinant promoter. In some embodiments, the promoter is a promoter capable of expressing the guide RNA in a mammalian cell. In some embodiments, the promoter is a promoter capable of expressing the guide RNA in a human cell. In some embodiments, the promoter is a promoter capable of expressing the guide RNA and restricting the guide RNA to the nucleus of the cell. In some embodiments, the promoter is a human RNA polymerase promoter or a sequence isolated or derived from a sequence encoding a human RNA polymerase promoter. In some embodiments, the promoter is a U6 promoter or a sequence isolated or derived from a sequence encoding a U6 promoter. In some embodiments, the U6 promoter is a human U6 promoter. In some embodiments, the promoter is a human tRNA promoter or a sequence isolated or derived from a sequence encoding a human tRNA promoter. In some embodiments, the promoter is a human valine tRNA promoter or a sequence isolated or derived from a sequence encoding a human valine tRNA promoter.

**[0512]** In some embodiments of the compositions of the disclosure, a promoter to drive expression of the guide RNA further comprises a regulatory element. In some embodiments, a vector comprising a promoter sequence to drive expression of the guide RNA further comprises a regulatory element. In some embodiments, a regulatory element enhances expression of the guide RNA. Exemplary regulatory elements include, but are not limited to, an enhancer element, an intron, an exon, or a combination thereof.

**[0513]** In some embodiments of the compositions of the disclosure, a vector of the disclosure comprises one or more of a sequence encoding a guide RNA, a promoter sequence to drive expression of the guide RNA and a sequence encoding a regulatory element. In some embodiments of the compositions of the disclosure, the vector further comprises a sequence encoding a fusion protein of the disclosure.

[0514] In some embodiments, a promoter driving expression of one or more gRNA of the disclosure is a human U6 promoter. In some aspects, a nucleic acid sequence encoding a human U6 promoter is set forth in SEQ ID NO: 448.

[0515] In some embodiments, a promoter driving expression of one or more RNA-binding polypeptides of the disclosure is an EFS promoter. In some aspects, a nucleic acid sequence encoding an EFS promoter is set forth in SEQ ID NO: 451.

[0516] In some embodiments, a promoter driving expression of one or more RNA-binding polypeptides of the disclosure is an EFS-UBB promoter. In some aspects, a nucleic acid sequence encoding an EFS-UBB promoter is set forth in SEQ ID NO: 450. In some aspects, a nucleic acid sequence encoding an EFS-UBB promoter is set forth in SEQ ID NO: 453.

[0517] In some embodiments, a promoter driving expression of one or more RNA-binding polypeptides of the disclosure is a CMV promoter. In some aspects, a nucleic acid sequence encoding a CMV promoter is set forth in SEQ ID NO: 12277.

[0518] In some embodiments, a nucleic acid sequence encoding a polyA sequence is set forth in SEQ ID NO: 457.

[0519] In some embodiments, the vector is a viral vector. In some embodiments, the vector is an adenoviral vector, an adeno-associated viral (AAV) vector, or a lentiviral vector. In some embodiments, the vector is a retroviral vector, an adenoviral/retroviral chimera vector, a herpes simplex viral I or II vector, a parvoviral vector, a reticuloendotheliosis viral vector, a polioviral vector, a papillomaviral vector, a vaccinia viral vector, or any hybrid or chimeric vector incorporating favorable aspects of two or more viral vectors. In some embodiments, the vector further comprises one or more expression control elements operably linked to the polynucleotide. In some embodiments, the vector further comprises one or more selectable markers. In some embodiments, the AAV vector has low toxicity. In some embodiments, the AAV vector does not incorporate into the host genome, thereby having a low probability of causing insertional mutagenesis. In some embodiments, the AAV vector can encode a range of total polynucleotides from 4.5 kb to 4.75 kb. In some embodiments, exemplary AAV vectors that may be used in any of the herein described compositions, systems, methods, and kits can include an AAV1 vector, a modified AAV1 vector, an AAV2 vector, a modified AAV2 vector, an AAV2-Tyr mutant vector, an AAV3 vector, a modified AAV3 vector, an AAV4 vector, a modified AAV4 vector, an AAV5 vector, a modified AAV5 vector, an AAV6 vector, a modified AAV6 vector, an AAV7 vector, a modified AAV7 vector, an AAV8 vector, an AAV9 vector, an AAV.rh10 vector, a modified AAV.rh10 vector, an AAVrh.74, an AAV.rh32/33 vector, a modified AAV.rh32/33 vector, an

AAV.rh43 vector, a modified AAV.rh43 vector, an AAV.rh64R1 vector, and a modified AAV.rh64R1 vector, an AAV-Tyr mutant vector, and any combinations or equivalents thereof. In some embodiments, the lentiviral vector is an integrase-competent lentiviral vector (ICLV). In some embodiments, the lentiviral vector can refer to the transgene plasmid vector as well as the transgene plasmid vector in conjunction with related plasmids (e.g., a packaging plasmid, a rev expressing plasmid, an envelope plasmid) as well as a lentiviral-based particle capable of introducing exogenous nucleic acid into a cell through a viral or viral-like entry mechanism. Lentiviral vectors are well-known in the art (see, e.g., Trono D. (2002) *Lentiviral vectors*, New York: Springer-Verlag Berlin Heidelberg and Durand et al. (2011) *Viruses* 3(2):132-159 doi: 10.3390/v3020132). In some embodiments, exemplary lentiviral vectors that may be used in any of the herein described compositions, systems, methods, and kits can include a human immunodeficiency virus (HIV) 1 vector, a modified human immunodeficiency virus (HIV) 1 vector, a human immunodeficiency virus (HIV) 2 vector, a modified human immunodeficiency virus (HIV) 2 vector, a sooty mangabey simian immunodeficiency virus (SIV<sub>SM</sub>) vector, a modified sooty mangabey simian immunodeficiency virus (SIV<sub>SM</sub>) vector, a African green monkey simian immunodeficiency virus (SIV<sub>AGM</sub>) vector, a modified African green monkey simian immunodeficiency virus (SIV<sub>AGM</sub>) vector, an equine infectious anemia virus (EIAV) vector, a modified equine infectious anemia virus (EIAV) vector, a feline immunodeficiency virus (FIV) vector, a modified feline immunodeficiency virus (FIV) vector, a Visna/maedi virus (VNV/VMV) vector, a modified Visna/maedi virus (VNV/VMV) vector, a caprine arthritis-encephalitis virus (CAEV) vector, a modified caprine arthritis-encephalitis virus (CAEV) vector, a bovine immunodeficiency virus (BIV), or a modified bovine immunodeficiency virus (BIV).

**[0520]** The disclosure provides adeno-associated virus (AAV) vectors comprising RNA-binding polypeptides capable of binding pre-mRNA sequences. In some aspects, AAV vectors of the disclosure comprise Cas13d constructs capable of binding target RNA sequences disclosed herein.

**[0521]** In some embodiments, murine USH2A exon 12 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 7. In some embodiments, the sequence encoding the vector of Table 7 comprises SEQ ID NO: 508.

**Table 7: AAV Vector A01270 targeting murine USH2A exon 12**

| Vector Element                                                                | Nucleotide sequence |
|-------------------------------------------------------------------------------|---------------------|
| Human U6 promoter                                                             | SEQ ID NO: 448      |
| Seq212 DR                                                                     | SEQ ID NO: 12279    |
| Spacer (mg1)                                                                  | SEQ ID NO: 413      |
| Termination                                                                   | TTTTTT              |
| Human U6 promoter                                                             | SEQ ID NO: 448      |
| Seq212 DR                                                                     | SEQ ID NO: 12279    |
| Spacer (mg7)                                                                  | SEQ ID NO: 418      |
| termination                                                                   | TTTTTT              |
| EFS                                                                           | SEQ ID NO: 451      |
| Kozak                                                                         | SEQ ID NO: 452      |
| dSeq212                                                                       | SEQ ID NO: 395      |
| Linker                                                                        | GGAAGC              |
| SV-40 NLS                                                                     | SEQ ID NO: 449      |
| Linker                                                                        | GAGGAC              |
| HA tag                                                                        | SEQ ID NO: 455      |
| Stop codon                                                                    | taa                 |
| Woodchuck Hepatitis Virus (WHP) Posttranscriptional Regulatory Element (WPRE) | SEQ ID NO: 456      |
| polyA signal                                                                  | SEQ ID NO: 457      |

**[0522]** In some embodiments, murine USH2A exon 12 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 8. In some embodiments, the sequence encoding the vector of Table 8 comprises SEQ ID NO: 509.

**Table 8: AAV Vector A01271 targeting murine USH2A exon 12**

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (mg1)      | SEQ ID NO: 413      |
| Termination       | TTTTTT              |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |

|                   |                |
|-------------------|----------------|
| Spacer (mg7)      | SEQ ID NO: 418 |
| termination       | TTTTTT         |
| EFS-UBB           | SEQ ID NO: 453 |
| Kozak             | SEQ ID NO: 452 |
| dSeq212           | SEQ ID NO: 395 |
| Linker            | GGAAGC         |
| SV-40 NLS         | SEQ ID NO: 449 |
| Linker            | GAGGAC         |
| HA tag            | SEQ ID NO: 455 |
| Stop codon        | taa            |
| WPRE              | SEQ ID NO: 456 |
| SV40 polyA signal | SEQ ID NO: 457 |

**[0523]** In some embodiments, murine USH2A exon 12 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 9. In some embodiments, the sequence encoding the vector of Table 9 comprises SEQ ID NO: 510.

Table 9: AAV Vector A01285 targeting murine USH2A exon 12

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 213      |
| Spacer (mg1-30nt) | SEQ ID NO: 414      |
| Seq212 DR         | SEQ ID NO: 213      |
| Spacer (mg7-30nt) | SEQ ID NO: 419      |
| termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| SV40 polyA signal | SEQ ID NO: 457      |

**[0524]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 10.

In some embodiments, the sequence encoding the vector of Table 10 comprises SEQ ID NO: 511.

Table 10: AAV Vector A01292 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g7)       | SEQ ID NO: 409      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |

**[0525]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 11. In some embodiments, the sequence encoding the vector of Table 11 comprises SEQ ID NO: 512.

Table 11: AAV Vector A01293 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g7)       | SEQ ID NO: 409      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |

|              |                |
|--------------|----------------|
| polyA signal | SEQ ID NO: 457 |
|--------------|----------------|

[0526] In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 12. In some embodiments, the sequence encoding the vector of Table 12 comprises SEQ ID NO: 513.

Table 12: AAV Vector A01296 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |

[0527] In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 13. In some embodiments, the sequence encoding the vector of Table 13 comprises SEQ ID NO: 514.

Table 13: AAV Vector A01297 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |

|              |                |
|--------------|----------------|
| HA tag       | SEQ ID NO: 455 |
| Stop codon   | taa            |
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |

**[0528]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 14. In some embodiments, the sequence encoding the vector of Table 14 comprises SEQ ID NO: 490.

Table 14: AAV Vector A01361 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g7)       | SEQ ID NO: 409      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0529]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 15. In some embodiments, the sequence encoding the vector of Table 15 comprises SEQ ID NO: 491.

Table 15: AAV Vector A01362 targeting human USH2A exon 13

| Vector Element | Nucleotide sequence |
|----------------|---------------------|
| 5' ITR         | SEQ ID NO: 458      |

|                   |                |
|-------------------|----------------|
| Human U6 promoter | SEQ ID NO: 448 |
| Seq42 DR          | SEQ ID NO: 462 |
| Spacer (g7)       | SEQ ID NO: 409 |
| Termination       | TTTTTT         |
| EFS-UBB           | SEQ ID NO: 453 |
| Kozak             | SEQ ID NO: 452 |
| dSeq212           | SEQ ID NO: 395 |
| Linker            | GGAAGC         |
| SV-40 NLS         | SEQ ID NO: 460 |
| Linker            | GAGGAC         |
| HA tag            | SEQ ID NO: 455 |
| Stop codon        | taa            |
| WPRE              | SEQ ID NO: 456 |
| polyA signal      | SEQ ID NO: 457 |
| 3' ITR            | SEQ ID NO: 459 |

**[0530]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 16. In some embodiments, the sequence encoding the vector of Table 16 comprises SEQ ID NO: 492.

Table 16: AAV Vector A01363 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0531]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 17. In some embodiments, the sequence encoding the vector of Table 17 comprises SEQ ID NO: 493.

Table 17: AAV Vector A01364 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0532]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 18. In some embodiments, the sequence encoding the vector of Table 18 comprises SEQ ID NO: 494.

Table 18: AAV Vector A01365 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |

|              |                |
|--------------|----------------|
| Kozak        | SEQ ID NO: 452 |
| dSeq212      | SEQ ID NO: 395 |
| Linker       | GGAAGC         |
| SV-40 NLS    | SEQ ID NO: 449 |
| Linker       | GAGGAC         |
| HA tag       | SEQ ID NO: 455 |
| Stop codon   | taa            |
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

**[0533]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 19. In some embodiments, the sequence encoding the vector of Table 19 comprises SEQ ID NO: 495.

Table 19: AAV Vector A01366 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0534]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 20.

In some embodiments, the sequence encoding the vector of Table 20 comprises SEQ ID NO: 496.

Table 20: AAV Vector A01367 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

[0535] In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 21. In some embodiments, the sequence encoding the vector of Table 21 comprises SEQ ID NO: 497.

Table 21: AAV Vector A01368 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq42 DR          | SEQ ID NO: 462      |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |

|              |                |
|--------------|----------------|
| Kozak        | SEQ ID NO: 452 |
| dSeq42       | SEQ ID NO: 461 |
| Linker       | GGAAGC         |
| SV-40 NLS    | SEQ ID NO: 460 |
| Linker       | GAGGAC         |
| HA tag       | SEQ ID NO: 455 |
| Stop codon   | taa            |
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

**[0536]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 22. In some embodiments, the sequence encoding the vector of Table 22 comprises SEQ ID NO: 498.

Table 22: AAV Vector A01369 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | (SEQ ID NO: 458)    |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq42 DR          | SEQ ID NO: 462      |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dSeq42            | SEQ ID NO: 461      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 460      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0537]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 23. In some embodiments, the sequence encoding the vector of Table 23 comprises SEQ ID NO: 499.

Table 23: AAV Vector A01370 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq42 DR          | SEQ ID NO: 462      |
| Spacer (g7)       | SEQ ID NO: 409      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dSeq42            | SEQ ID NO: 461      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 460      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0538]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 24. In some embodiments, the sequence encoding the vector of Table 24 comprises SEQ ID NO: 500.

Table 24: AAV Vector A01371 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq42 DR          | SEQ ID NO: 462      |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |

|              |                |
|--------------|----------------|
| Kozak        | SEQ ID NO: 452 |
| dSeq42       | SEQ ID NO: 461 |
| Linker       | GGAAGC         |
| SV-40 NLS    | SEQ ID NO: 460 |
| Linker       | GAGGAC         |
| HA tag       | SEQ ID NO: 455 |
| Stop codon   | Taa            |
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

**[0539]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 25. In some embodiments, the sequence encoding the vector of Table 25 comprises SEQ ID NO: 501.

Table 25: AAV Vector A01372 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq42 DR          | SEQ ID NO: 462      |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dSeq42            | SEQ ID NO: 461      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 460      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0540]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 26.

In some embodiments, the sequence encoding the vector of Table 26 comprises SEQ ID NO: 502.

Table 26: AAV Vector A01373 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| CasRx DR          | SEQ ID NO: 302      |
| Spacer (g7)       | SEQ ID NO: 409      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dCasRx            | SEQ ID NO: 463      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

[0541] In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 27. In some embodiments, the sequence encoding the vector of Table 27 comprises SEQ ID NO: 503.

Table 27: AAV Vector A01374 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| CasRx DR          | SEQ ID NO: 302      |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |

|              |                |
|--------------|----------------|
| dCasRx       | SEQ ID NO: 463 |
| Linker       | GGAAGC         |
| SV-40 NLS    | SEQ ID NO: 449 |
| Linker       | GAGGAC         |
| HA tag       | SEQ ID NO: 455 |
| Stop codon   | taa            |
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

**[0542]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 28. In some embodiments, the sequence encoding the vector of Table 28 comprises SEQ ID NO: 504.

Table 28: AAV Vector A01375 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| CasRx DR          | SEQ ID NO: 302      |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS-UBB           | SEQ ID NO: 453      |
| Kozak             | SEQ ID NO: 452      |
| dCasRx            | SEQ ID NO: 463      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0543]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 29. In some embodiments, the sequence encoding the vector of Table 29 comprises SEQ ID NO: 505.

Table 29: AAV Vector A01376 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| CasRx DR          | SEQ ID NO: 302      |
| Spacer (g7)       | SEQ ID NO: 409      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dCasRx            | SEQ ID NO: 463      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0544]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 30. In some embodiments, the sequence encoding the vector of Table 30 comprises SEQ ID NO: 506.

Table 30: AAV Vector A01377 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| CasRx DR          | SEQ ID NO: 302      |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dCasRx            | SEQ ID NO: 463      |
| Linker            | GGAAGC              |

|              |                |
|--------------|----------------|
| SV-40 NLS    | SEQ ID NO: 449 |
| Linker       | GAGGAC         |
| HA tag       | SEQ ID NO: 455 |
| Stop codon   | taa            |
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

**[0545]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 31. In some embodiments, the sequence encoding the vector of Table 31 comprises SEQ ID NO: 507.

Table 31: AAV Vector A01378 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| CasRx DR          | SEQ ID NO: 302      |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dCasRx            | SEQ ID NO: 463      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0546]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 32. In some embodiments, the sequence encoding the vector of Table 32 comprises SEQ ID NO: 515.

Table 32: AAV Vector A01493 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g7)       | SEQ ID NO: 409      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

[0547] In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 33. In some embodiments, the sequence encoding the vector of Table 33 comprises SEQ ID NO: 516.

Table 33: AAV Vector A01494 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |

|              |                |
|--------------|----------------|
| HA tag       | SEQ ID NO: 455 |
| Stop codon   | taa            |
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

**[0548]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 34. In some embodiments, the sequence encoding the vector of Table 34 comprises SEQ ID NO: 517.

Table 34: AAV Vector A01495 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0549]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 35. In some embodiments, the sequence encoding the vector of Table 35 comprises SEQ ID NO: 518.

Table 35: AAV Vector A01496 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq42 DR          | SEQ ID NO: 462      |
| Spacer (g7)       | SEQ ID NO: 409      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dSeq42            | SEQ ID NO: 461      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 460      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0550]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 36. In some embodiments, the sequence encoding the vector of Table 36 comprises SEQ ID NO: 519.

Table 36: AAV Vector A01497 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq42 DR          | SEQ ID NO: 462      |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dSeq42            | SEQ ID NO: 461      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 460      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |

|              |                |
|--------------|----------------|
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

**[0551]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 37. In some embodiments, the sequence encoding the vector of Table 37 comprises SEQ ID NO: 520.

Table 37: AAV Vector A01498 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq42 DR          | SEQ ID NO: 462      |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dSeq42            | SEQ ID NO: 461      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 460      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0552]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 38. In some embodiments, the sequence encoding the vector of Table 38 comprises SEQ ID NO: 521.

Table 38: AAV Vector A01499 targeting human USH2A exon 13

| Vector Element | Nucleotide sequence |
|----------------|---------------------|
| 5' ITR         | SEQ ID NO: 458      |

|                   |                |
|-------------------|----------------|
| Human U6 promoter | SEQ ID NO: 448 |
| CasRx DR          | SEQ ID NO: 302 |
| Spacer (g7)       | SEQ ID NO: 409 |
| Termination       | TTTTTT         |
| EFS-UBB-1         | SEQ ID NO: 450 |
| Kozak             | SEQ ID NO: 452 |
| dCasRx            | SEQ ID NO: 463 |
| Linker            | GGAAGC         |
| SV-40 NLS         | SEQ ID NO: 44  |
| Linker            | GAGGAC         |
| HA tag            | SEQ ID NO: 455 |
| Stop codon        | taa            |
| WPRE              | SEQ ID NO: 456 |
| polyA signal      | SEQ ID NO: 457 |
| 3' ITR            | SEQ ID NO: 459 |

**[0553]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 39. In some embodiments, the sequence encoding the vector of Table 39 comprises SEQ ID NO: 522.

Table 39: AAV Vector A01500 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| CasRx DR          | SEQ ID NO: 302      |
| Spacer (g9)       | SEQ ID NO: 411      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dCasRx            | SEQ ID NO: 463      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |

|              |                |
|--------------|----------------|
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

**[0554]** In some embodiments, human USH2A exon 13 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 40. In some embodiments, the sequence encoding the vector of Table 40 comprises SEQ ID NO: 523.

Table 40: AAV Vector A01501 targeting human USH2A exon 13

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| CasRx DR          | SEQ ID NO: 302      |
| Spacer (g10)      | SEQ ID NO: 412      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dCasRx            | SEQ ID NO: 463      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

**[0555]** In some embodiments, human DMD exon 51 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 41. In some embodiments, the sequence encoding the vector of Table 41 comprises SEQ ID NO: 524.

Table 41: AAV Vector A01511 targeting human DMD exon 51

| Vector Element | Nucleotide sequence |
|----------------|---------------------|
|----------------|---------------------|

|                   |                  |
|-------------------|------------------|
| 5' ITR            | SEQ ID NO: 458   |
| Human U6 promoter | SEQ ID NO: 448   |
| Seq212 DR         | SEQ ID NO: 12279 |
| Spacer (g5)       | SEQ ID NO: 426   |
| Termination       | TTTTTT           |
| EFS               | SEQ ID NO: 451   |
| Kozak             | SEQ ID NO: 452   |
| dSeq212           | SEQ ID NO: 395   |
| Linker            | GGAAGC           |
| SV-40 NLS         | SEQ ID NO: 449   |
| Linker            | GAGGAC           |
| HA tag            | SEQ ID NO: 455   |
| Stop codon        | taa              |
| WPRE              | SEQ ID NO: 456   |
| polyA signal      | SEQ ID NO: 457   |
| 3' ITR            | SEQ ID NO: 459   |

[0556] In some embodiments, human DMD exon 45 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 42. In some embodiments, the sequence encoding the vector of Table 42 comprises SEQ ID NO: 525.

Table 42: AAV Vector A01512 targeting human DMD exon 45

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (gexon45)  | SEQ ID NO: 436      |
| Termination       | TTTTTT              |
| EFS               | SEQ ID NO: 451      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |

|        |                |
|--------|----------------|
| 3' ITR | SEQ ID NO: 459 |
|--------|----------------|

[0557] In some embodiments, human DMD exon 51 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 43. In some embodiments, the sequence encoding the vector of Table 43 comprises SEQ ID NO: 526.

Table 43: AAV Vector A01513 targeting human DMD exon 51

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (g5)       | SEQ ID NO: 426      |
| Termination       | TTTTTT              |
| EFS-UBB-1         | SEQ ID NO: 450      |
| Kozak             | SEQ ID NO: 452      |
| dSeq212           | SEQ ID NO: 395      |
| Linker            | GGAAGC              |
| SV-40 NLS         | SEQ ID NO: 449      |
| Linker            | GAGGAC              |
| HA tag            | SEQ ID NO: 455      |
| Stop codon        | taa                 |
| WPRE              | SEQ ID NO: 456      |
| polyA signal      | SEQ ID NO: 457      |
| 3' ITR            | SEQ ID NO: 459      |

[0558] In some embodiments, human DMD exon 45 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table 44. In some embodiments, the sequence encoding the vector of Table 44 comprises SEQ ID NO: 527.

Table 44: AAV Vector A01514 targeting human DMD exon 45

| Vector Element    | Nucleotide sequence |
|-------------------|---------------------|
| 5' ITR            | SEQ ID NO: 458      |
| Human U6 promoter | SEQ ID NO: 448      |
| Seq212 DR         | SEQ ID NO: 12279    |
| Spacer (gexon45)  | SEQ ID NO: 436      |
| Termination       | TTTTTT              |

|              |                |
|--------------|----------------|
| EFS-UBB-1    | SEQ ID NO: 450 |
| Kozak        | SEQ ID NO: 452 |
| dSeq212      | SEQ ID NO: 395 |
| Linker       | GGAAGC         |
| SV-40 NLS    | SEQ ID NO: 449 |
| Linker       | GAGGAC         |
| HA tag       | SEQ ID NO: 455 |
| Stop codon   | taa            |
| WPRE         | SEQ ID NO: 456 |
| polyA signal | SEQ ID NO: 457 |
| 3' ITR       | SEQ ID NO: 459 |

[0559] In some embodiments, human DMD exon 51 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AA. In some embodiments, the sequence encoding the vector of Table AA comprises SEQ ID NO: 12280.

Table AA: AAV Vector A02636 targeting human DMD exon 51

|                      |                  |
|----------------------|------------------|
| ITR                  | SEQ ID NO: 458   |
| hU6 promoter         | SEQ ID NO: 448   |
| DR                   | SEQ ID NO: 213   |
| spacer               | SEQ ID NO: 12275 |
| Termination sequence | TTTTTT           |
| CMV promoter         | SEQ ID NO: 12277 |
| Kozak                | SEQ ID NO: 452   |
| dSeq212-4aa          | SEQ ID NO: 12278 |
| Linker               | GGAAGC           |
| NLS                  | SEQ ID NO: 449   |
| Linker               | GAGGAC           |
| HA tag               | SEQ ID NO: 455   |
| Stop codon           | taa              |
| WPRE3                | SEQ ID NO: 456   |
| SV40 pA              | SEQ ID NO: 457   |
| ITR                  | SEQ ID NO: 459   |

[0560] In some embodiments, human DMD exon 51 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AB. In some embodiments, the sequence encoding the vector of Table AB comprises SEQ ID NO: 12281.

Table AB: AAV Vector A02637 targeting human DMD exon 51

|              |                |
|--------------|----------------|
| ITR          | SEQ ID NO: 458 |
| hU6 promoter | SEQ ID NO: 448 |
| DR           | SEQ ID NO: 213 |

|                      |                  |
|----------------------|------------------|
| spacer               | SEQ ID NO: 12276 |
| Termination sequence | TTTTTT           |
| CMV promoter         | SEQ ID NO: 12277 |
| Kozak                | SEQ ID NO: 452   |
| dSeq212-4aa          | SEQ ID NO: 12278 |
| Linker               | GGAAGC           |
| NLS                  | SEQ ID NO: 449   |
| Linker               | GAGGAC           |
| HA tag               | SEQ ID NO: 455   |
| Stop codon           | taa              |
| WPRES                | SEQ ID NO: 456   |
| SV40 pA              | SEQ ID NO: 457   |
| ITR                  | SEQ ID NO: 459   |

[0561] In some embodiments, human DMD exon 51 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AC. In some embodiments, the sequence encoding the vector of Table AC comprises SEQ ID NO: 12282.

Table AC: AAV Vector A02638 targeting human DMD exon 51

|                      |                  |
|----------------------|------------------|
| ITR                  | SEQ ID NO: 458   |
| hU6 promoter         | SEQ ID NO: 448   |
| DR                   | SEQ ID NO: 213   |
| spacer               | SEQ ID NO: 426   |
| Termination sequence | TTTTTT           |
| CMV promoter         | SEQ ID NO: 12277 |
| Kozak                | SEQ ID NO: 452   |
| dSeq212-4aa          | SEQ ID NO: 12278 |
| Linker               | GGAAGC           |
| NLS                  | SEQ ID NO: 449   |
| Linker               | GAGGAC           |
| HA tag               | SEQ ID NO: 455   |
| Stop codon           | taa              |
| WPRES                | SEQ ID NO: 456   |
| SV40 pA              | SEQ ID NO: 457   |
| ITR                  | SEQ ID NO: 459   |

[0562] In some embodiments, human DMD exon 51 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AD. In some embodiments, the sequence encoding the vector of Table AD comprises SEQ ID NO: 12283.

Table AD: AAV Vector A02639 targeting human DMD exon 51

|                      |                  |
|----------------------|------------------|
| ITR                  | SEQ ID NO: 458   |
| hU6 promoter         | SEQ ID NO: 448   |
| DR                   | SEQ ID NO: 12279 |
| spacer               | SEQ ID NO: 2307  |
| Termination sequence | TTTTTT           |
| CMV promoter         | SEQ ID NO: 12277 |

|             |                  |
|-------------|------------------|
| Kozak       | SEQ ID NO: 452   |
| dSeq212-4aa | SEQ ID NO: 12278 |
| Linker      | GGAAGC           |
| NLS         | SEQ ID NO: 449   |
| Linker      | GAGGAC           |
| HA tag      | SEQ ID NO: 455   |
| Stop codon  | taa              |
| WPRES       | SEQ ID NO: 456   |
| SV40 pA     | SEQ ID NO: 457   |
| ITR         | SEQ ID NO: 459   |

**[0563]** In some embodiments, human DMD exon 51 RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AE. In some embodiments, the sequence encoding the vector of Table AE comprises SEQ ID NO: 12284.

Table AE: AAV Vector A02640 targeting human DMD exon 51

|                      |                  |
|----------------------|------------------|
| ITR                  | SEQ ID NO: 458   |
| hU6 promoter         | SEQ ID NO: 448   |
| DR                   | SEQ ID NO: 12279 |
| spacer               | SEQ ID NO: 2247  |
| Termination sequence | TTTTTT           |
| CMV promoter         | SEQ ID NO: 12277 |
| Kozak                | SEQ ID NO: 452   |
| dSeq212-4aa          | SEQ ID NO: 12278 |
| Linker               | GGAAGC           |
| NLS                  | SEQ ID NO: 449   |
| Linker               | GAGGAC           |
| HA tag               | SEQ ID NO: 455   |
| Stop codon           | taa              |
| WPRES                | SEQ ID NO: 456   |
| SV40 pA              | SEQ ID NO: 457   |
| ITR                  | SEQ ID NO: 459   |

**[0564]** In some embodiments, human DMD RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AF. In some embodiments, the sequence encoding the vector of Table AF comprises SEQ ID NO: 12286.

Table AF: AAV Vector A03278 targeting human DMD exon 51 exon splicing enhancers

|                              |                  |
|------------------------------|------------------|
| hU6 promoter                 | SEQ ID NO: 44    |
| Seq212 DR                    | SEQ ID NO: 213   |
| g38+                         | SEQ ID NO: 12285 |
| Seq212 DR                    | SEQ ID NO: 213   |
| G42                          | SEQ ID NO: 12276 |
| CMV enhancer and promoter    | SEQ ID NO: 12277 |
| 5' untranslated region (UTR) | SEQ ID NO: 12314 |
| dSeq212/                     | SEQ ID NO: 12278 |
| 4aa mut                      |                  |
| SV40-NLS                     | SEQ ID NO: 449   |
| HA tag                       | SEQ ID NO: 455   |
| WPRES                        | SEQ ID NO: 456   |

|             |                |
|-------------|----------------|
| SV40 poly A | SEQ ID NO: 459 |
|-------------|----------------|

**[0565]** In some embodiments, human DMD RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AG. In some embodiments, the sequence encoding the vector of Table AG comprises SEQ ID NO: 12310.

Table AG: AAV Vector A03275 targeting human DMD exon 51 exon splicing enhancers

|                           |                  |
|---------------------------|------------------|
| hU6 promoter              | SEQ ID NO: 448   |
| Seq212 DR                 | SEQ ID NO: 213   |
| g5+                       | SEQ ID NO: 12311 |
| Seq212 DR                 | SEQ ID NO: 213   |
| g17+                      | SEQ ID NO: 12312 |
| CMV enhancer and promoter | SEQ ID NO: 12277 |
| 5'UTR                     | SEQ ID NO: 12314 |
| dSeq212/                  | SEQ ID NO: 12278 |
| SV40 NLS                  | SEQ ID NO: 449   |
| HA Tag                    | SEQ ID NO: 455   |
| WPRES                     | SEQ ID NO: 456   |
| SV40 poly A               | SEQ ID NO: 459   |

**[0566]** In some embodiments, human DMD RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AH. In some embodiments, the sequence encoding the vector of Table AH comprises SEQ ID NO: 12313.

Table AH: AAV Vector A03276 targeting human DMD exon 51 exon splicing enhancers

|                           |                  |
|---------------------------|------------------|
| hU6 promoter              | SEQ ID NO: 448   |
| Seq212 DR                 | SEQ ID NO: 213   |
| g5+                       | SEQ ID NO: 12311 |
| Seq212 DR                 | SEQ ID NO: 213   |
| g41                       | SEQ ID NO: 12275 |
| CMV enhancer and promoter | ID NO: 12277     |
| 5'UTR                     | SEQ ID NO: 12314 |
| dSeq212/                  | SEQ ID NO: 12278 |
| SV40 NLS                  | SEQ ID NO: 449   |
| HA Tag                    | SEQ ID NO: 455   |
| WPRES                     | SEQ ID NO: 456   |
| SV40 poly A               | SEQ ID NO: 459   |

**[0567]** In some embodiments, human DMD RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AI. In some embodiments, the sequence encoding the vector of Table AI comprises SEQ ID NO: 12315.

Table AI: AAV Vector A03277 targeting human DMD exon 51 exon splicing enhancers

|                           |                  |
|---------------------------|------------------|
| hU6 promoter              | SEQ ID NO: 448   |
| Seq212 DR                 | SEQ ID NO: 213   |
| g5+                       | SEQ ID NO: 12311 |
| Seq212 DR                 | SEQ ID NO: 213   |
| g42                       | SEQ ID NO: 12276 |
| CMV enhancer and promoter | SEQ ID NO: 12277 |
| 5'UTR                     | SEQ ID NO: 12314 |
| dSeq212/<br>4aa mut       | SEQ ID NO: 12278 |
| SV40 NLS                  | SEQ ID NO: 449   |
| HA Tag                    | SEQ ID NO: 455   |

|             |                |
|-------------|----------------|
| WPRES       | SEQ ID NO: 456 |
| SV40 poly A | SEQ ID NO: 459 |

**[0568]** In some embodiments, human DMD RNA-targeting Cas13d vectors comprise from 5' to 3' a nucleic acid sequence comprising the elements set forth in Table AJ. In some embodiments, the sequence encoding the vector of Table AJ comprises SEQ ID NO: 12316.

Table AJ: AAV Vector A03853 targeting human DMD exon 51 exon splicing enhancers

|                           |                  |
|---------------------------|------------------|
| hU6 promoter              | SEQ ID NO: 448   |
| Seq212 DR                 | SEQ ID NO: 213   |
| g38+                      | SEQ ID NO: 12285 |
| Seq212 DR                 | SEQ ID NO: 213   |
| g42                       | SEQ ID NO: 12276 |
| CMV enhancer and promoter | SEQ ID NO: 12277 |
| 5'UTR                     | SEQ ID NO: 12314 |
| dSeq212/<br>4aa mut       | SEQ ID NO: 12278 |
| WPRES                     | SEQ ID NO: 456   |
| SV40 poly A               | SEQ ID NO: 459   |

**[0569]** In some embodiments, nucleic acid sequences encoding RNA-targeting Cas13d compositions of the disclosure are codon optimized nucleic acid sequences. In some embodiments, the codon optimized sequence encoding an RNA-targeting Cas13d composition exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased translation in a human subject relative to a wild-type or non-codon optimized nucleic acid sequence.

**[0570]** In some aspects, a codon optimized nucleic acid sequence encoding an RNA-targeting Cas13d composition exhibits increased stability. In some aspects, a codon optimized nucleic acid sequence encoding an RNA-targeting Cas13d composition exhibits increased stability through increased resistance to hydrolysis. In some embodiments, the codon optimized sequence encoding an RNA-targeting Cas13d composition exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased stability relative to a wild-type or non-codon optimized nucleic acid sequence. In some embodiments, the codon optimized sequence encoding an RNA-targeting Cas13d composition exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased resistance to hydrolysis in a human subject relative to a wild-type or non-codon optimized nucleic acid sequence.

**[0571]** In some aspects, a codon optimized nucleic acid sequence encoding an RNA-targeting Cas13d composition can comprise no donor splice sites. In some aspects, a codon optimized nucleic acid sequence encoding an RNA-targeting dCas13d protein can comprise

no more than about one, or about two, or about three, or about four, or about five, or about six, or about seven, or about eight, or about nine, or about ten donor splice sites. In some aspects, a codon optimized nucleic acid sequence encoding an RNA-targeting Cas13d composition comprises at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten fewer donor splice sites as compared to a non-codon optimized nucleic acid sequence encoding the RNA-targeting Cas13d composition.

**[0572]** Without wishing to be bound by theory, the removal of donor splice sites in the codon optimized nucleic acid sequence can unexpectedly and unpredictably increase expression of the RNA-targeting dCas13d protein *in vivo*, as cryptic splicing is prevented. Moreover, cryptic splicing may vary between different subjects, meaning that the expression level of the RNA-targeting dCas13d protein comprising donor splice sites may unpredictably vary between different subjects. Such unpredictability is unacceptable in the context of human therapy. Accordingly, the codon optimized nucleic acid sequences which lacks donor splice sites, unexpectedly and surprisingly allows for increased expression of the RNA-targeting dCas13d protein in human subjects and regularizes expression of the RNA-targeting dCas13d protein across different human subjects.

**[0573]** In some aspects, a codon optimized nucleic acid sequence encoding an RNA-targeting Cas13d composition can have a GC content that differs from the GC content of the non-codon optimized nucleic acid sequence encoding the RNA-targeting Cas13d composition. In some aspects, the GC content of a codon optimized nucleic acid sequence encoding a RNA-targeting Cas13d composition is more evenly distributed across the entire nucleic acid sequence, as compared to the non-codon optimized nucleic acid sequence encoding the RNA-targeting Cas13d protein.

**[0574]** Without wishing to be bound by theory, by more evenly distributing the GC content across the entire nucleic acid sequence, the codon optimized nucleic acid sequence exhibits a more uniform melting temperature (“*T<sub>m</sub>*”) across the length of the transcript. The uniformity of melting temperature results unexpectedly in increased expression of the codon optimized nucleic acid in a human subject, as transcription and/or translation of the nucleic acid sequence occurs with less stalling of the polymerase and/or ribosome.

**[0575]** In some aspects, a codon optimized nucleic acid sequence encoding an RNA-targeting Cas13d composition can have fewer repressive microRNA target binding sites as compared to the non-codon optimized nucleic acid sequence encoding the RNA-targeting Cas13d composition. In some aspects, a codon optimized nucleic acid sequence encoding an

RNA-targeting Cas13d composition can have at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten, or at least ten fewer repressive microRNA target binding sites as compared to the non-codon optimized nucleic acid sequence the RNA-targeting Cas13d composition.

[0576] Without wishing to be bound by theory, by having fewer repressive microRNA target binding sites, the codon optimized nucleic acid sequence encoding an RNA-targeting Cas13d composition unexpectedly exhibits increased expression in a human subject.

### **Fusion Proteins**

[0577] In some embodiments of the compositions and methods of the disclosure, the composition comprises a sequence encoding a target RNA-binding fusion protein comprising (a) a sequence encoding a first RNA-binding polypeptide or portion thereof; and optionally (b) a sequence encoding a polypeptide capable of modulating RNA splicing. In some embodiments of the compositions and methods of the disclosure, the composition comprises a sequence encoding a target RNA-binding fusion protein comprising (a) a sequence encoding a first RNA-binding polypeptide or portion thereof; and optionally (b) a sequence encoding a second RNA-binding polypeptide. In some embodiments, the first RNA-binding polypeptide binds a target pre-mRNA, and the second RNA-binding polypeptide comprises a polypeptide capable of modulating RNA splicing. In embodiments, the first RNA-binding polypeptide binds a target RNA, and the second RNA-binding polypeptide comprises RNA-nuclease activity.

[0578] In some aspects, polypeptides capable of modulating RNA splicing include, but are not limited to, components of the spliceosome, RNA Binding Fox-1 Homolog 1 (RBFOX1), RNA Binding Motif Protein 38 (RMP38), heterogeneous nuclear ribonucleoprotein A1 (hnRNPA1), and alternative splicing factor-1 (ASF1).

[0579] In some embodiments, a target RNA-binding fusion protein is an RNA-guided target RNA-binding fusion protein. RNA-guided target RNA-binding fusion proteins comprise at least one RNA-binding polypeptide which corresponds to a gRNA which guides the RNA-binding polypeptide to target RNA. RNA-guided target RNA-binding fusion proteins include without limitation, RNA-binding polypeptides which are CRISPR/Cas-based RNA-binding polypeptides or portions thereof.

### **Signal Sequences**

[0580] In some embodiments, a target RNA-binding protein of the disclosure comprises a signal sequence. In some embodiments, a target RNA-binding protein comprises one or more

signal sequences. In some embodiments, the signal sequence(s) is a nuclear localization sequence (NLS), nuclear export signal (NES) or a combination thereof. In some embodiments, the tag sequence comprises a nuclear localization sequence (NLS). In some embodiments, the NLS sequence comprises a sequence listed in table 45. In some embodiments, the NLS signal sequence is a human NLS. In some embodiments, the human NLS signal sequence is a human pRB-NLS or a human pRB-NLS (extended version).

Table 45: Nuclear Localization Sequences of the disclosure

| Name                                    | Amino acid Sequence  | SEQ ID NO: |
|-----------------------------------------|----------------------|------------|
| <b>SV40-NLS</b>                         | PKKKRKV              | 367        |
| <b>human H2B-NLS</b>                    | GKKRKRSRK            | 368        |
| <b>yeast H2B-NLS</b>                    | GKKRSKV              | 369        |
| <b>human p53-NLS</b>                    | KRALPNNTSSSPQPKKKP   | 370        |
| <b>human-cmyc-NLS</b>                   | PAAKRVKLD            | 371        |
| <b>human pRB-NLS</b>                    | KRSAEGSNPPKPLKKLR    | 372        |
| <b>human Nucleoplasmin-NLS</b>          | KRPAATKKAGQAKKKKLDK  | 373        |
| <b>Human pRB-NLS (extended version)</b> | DRVLRSAEGSNPPKPLKKLR | 374        |

[0581] In some embodiments, a nucleic acid sequence encoding an SV40 NLS is set forth in SEQ ID NO: 449. In some embodiments, a nucleic acid sequence encoding an SV40 NLS is set forth in SEQ ID NO: 460.

[0582] In some embodiments, the signal sequence comprises one or more NES sequences. In some embodiments, the one or more NES sequence comprises a sequence listed in Table 46.

Table 46: Nuclear Export Sequences of the disclosure

| Name                 | Amino acid Sequence | SEQ ID NO: |
|----------------------|---------------------|------------|
| <b>HIV REV NES</b>   | LPPLERLTLD          | 377        |
| <b>Human PKI NES</b> | LALKLAGLDI          | 378        |

[0583] In some embodiments, a target RNA-binding protein of the disclosure comprises a tag sequence. In some embodiments, the tag sequence is a FLAG tag.

[0584] In some embodiments, the FLAG tag sequence is DYKDDDDK (SEQ ID NO: 366).

#### Linker Sequences

[0585] In some embodiments, a target RNA-binding protein comprises a linker sequence. In some embodiments, the linker sequence may comprise or consist of 1, 2, 3, 4, 5, 6, 7, 8, 9,

10, 15, 20, 25, 30, 35, 40, 45, 50 or any number of amino acids in between. In some embodiments, the linker sequence comprises a linker sequence listed in Table 47.

Table 47. Linker Sequences of the disclosure

| <b>Linker Sequence (amino acid)</b> | <b>SEQ ID NO:</b> |
|-------------------------------------|-------------------|
| <b>GGS</b>                          | 340               |
| <b>VDTANGS</b>                      | 341               |
| <b>VDTGNGS</b>                      | 342               |
| <b>SGSETPGTSESATPES</b>             | 343               |
| <b>GGGGSGGGGS</b>                   | 344               |
| <b>GGGGSGGGGS</b>                   | 345               |
| <b>GGGGSGGGGS</b>                   | 346               |
| <b>GGS</b>                          | 347               |
| <b>EAAAKEAAAK</b>                   | 348               |
| <b>EAAAKEAAAKEAAAK</b>              | 349               |
| <b>EAAAKEAAAKEAAAKEAAAK</b>         | 350               |
| <b>APAPAPAP</b>                     | 351               |
| <b>APAPAPAPAPAP</b>                 | 352               |
| <b>APAPAPAPAPAPAPAPAP</b>           | 353               |
| <b>GGGGSEAAAK</b>                   | 354               |
| <b>EAAAKGGGGS</b>                   | 355               |
| <b>GGGGSGGGGSEAAAKEAAAK</b>         | 356               |
| <b>EAAAKEAAAKGGGGSGGGGS</b>         | 357               |
| <b>RQTSPDPCPQLPLVPR</b>             | 358               |
| <b>VDTGNWF</b>                      | 359               |
| <b>VDTANGSVDTGNGS</b>               | 360               |
| <b>ARNVEERLCL</b>                   | 361               |
| <b>AIELNPSNA</b>                    | 362               |
| <b>ICGSRNL</b>                      | 363               |
| <b>VLATDMSKH</b>                    | 364               |
| <b>FLRELPEP</b>                     | 365               |
| <b>LIPKDQYYC</b>                    | 365               |

### **Non-Guided RNA-Binding Proteins**

[0586] In some embodiments, a target RNA-binding protein is not an RNA-guided target RNA-binding protein and as such comprises at least one RNA-binding polypeptide which is capable of binding a target RNA without a corresponding gRNA sequence. Such non-guided RNA-binding polypeptides include, without limitation, at least one RNA-binding protein or RNA-binding portion thereof which is a PUF (Pumilio and FBF homology family) protein. This type RNA-binding polypeptide can be used instead of a gRNA-guided RNA binding protein such as CRISPR/Cas. The unique RNA recognition mode of PUF proteins (named

for *Drosophila* Pumilio and *C. elegans* fem-3 binding factor) that are involved in mediating mRNA stability and translation are well known in the art. The PUF domain of human Pumilio1, also known in the art, binds tightly to cognate RNA sequences and its specificity can be modified. It contains eight PUF modules that recognize eight consecutive RNA bases with each module recognizing a single base. Since two amino acid side chains in each module recognize the Watson-Crick edge of the corresponding base and determine the specificity of that module, a PUF protein can be designed to specifically bind most 8 to 16-nt RNA. Wang *et al.*, *Nat Methods*. 2009; 6(11): 825-830. See also WO2012/068627 which is incorporated by reference herein in its entirety.

**[0587]** The modular nature of the PUF-RNA interaction has been used to rationally engineer the binding specificity of PUF domains (Cheong, C. G. & Hall, T. M. (2006) PNAS 103: 13635-13639; Wang, X. et al (2002) Cell 110: 501-512). However, only the successful design of PUF proteins with modules that recognize adenine, guanine or uracil have been reported prior to the teachings of WO2012/06827 *supra*. While the wild-type PumHD does not bind cytosine (C), molecular engineering has shown that some of the Pum units can be mutated to bind C with good yield and specificity. See e.g., Dong, S. et al. Specific and modular binding code for cytosine recognition in Pumilio/FBF (PUF) RNA-binding domains, *The Journal of biological chemistry* 286, 26732-26742 (2011). Accordingly, PumHD is a modified version of the WT Pumilio protein that exhibits programmable binding to arbitrary 8-base sequences of RNA. Each of the eight units of PumHD can bind to all four RNA bases, and the RNA bases flanking the target sequence do not affect binding. See also the following for art-recognized RNA-binding rules of PUF design: Filipovska A, Razif MF, Nygård KK, & Rackham O. A universal code for RNA recognition by PUF proteins. *Nature chemical biology*, 7(7), 425-427 (2011); Filipovska A, & Rackham O. Modular recognition of nucleic acids by PUF, TALE and PPR proteins. *Molecular BioSystems*, 8(3), 699-708 (2012); Abil Z, Denard CA, & Zhao H. Modular assembly of designer PUF proteins for specific post-transcriptional regulation of endogenous RNA. *Journal of biological engineering*, 8(1), 7 (2014); Zhao Y, Mao M, Zhang W, Wang J, Li H, Yang Y, Wang Z, & Wu J. Expanding RNA binding specificity and affinity of engineered PUF domains. *Nucleic Acids Research*, 46(9), 4771-4782 (2018); Shinoda K, Tsuji S, Futaki S, & Imanishi M. Nested PUF Proteins: Extending Target RNA Elements for Gene Regulation. *ChemBioChem*, 19(2), 171-176 (2018); Koh YY, Wang Y, Qiu C, Opperman L, Gross L, Tanaka Hall TM, & Wickens M. Stacking Interactions in PUF-RNA Complexes. *RNA*, 17(4), 718-727 (2011).

**[0588]** As such, it is well known in the art that human PUM1 (1186 amino acids) contains an RNA-binding domain (RBD) in the C-terminus of the protein (also known as Pumilio homology domain PUM-HD amino acid 828-amino acid 1175) and that PUFs are based on the RBD of human PUM1. There are 8 structural repeat modules of 36 amino acids (except module 7 which has 43 amino acids) for RNA binding and flanking N- and C- terminal regions important for protein structure and stability. Within each repeat module, amino acids 12, 13, and 16 are important for RNA binding with 12 and 16 responsible for RNA base recognition. Amino acid 13 stacks with RNA bases and can be modified to tune specificity and affinity. Alternatively, the PUF design may maintain amino acid 13 as human PUM1's native residue. In some embodiments of the PUF or PUMBY compositions disclosed herein, amino acid 13 (for stacking) will be engineered with an H and in other embodiments, will be engineered with a Y. In some embodiments, stacking residues may be modified to improve binding and specificity. Recognition occurs in reverse orientation as N- to C-terminal PUF recognizes 3' to 5' RNA. Accordingly, PUF engineering of 8 modules (8PUF), as known in the art, mimics a human protein. An exemplary 8-mer RNA recognition (8PUF) would be designed as follows: R1'-R1-R2-R3-R4-R5-R6-R7-R8-R8'. In one embodiment, an 8PUF is used as the RBD. In another embodiment, a variation of the 8PUF design is used to create a 14-mer RNA recognition (14PUF) RBD, 15-mer RNA recognition (15PUF) RBD, or a 16-mer RNA recognition (16PUF) RBD. In another embodiment, the PUF can be engineered to comprise a 4-mer, 5-mer, 6-mer, 7-mer, 8-mer, 9-mer, 10-mer, 11-mer, 12-mer, 13-mer, 14-mer, 15-mer, 16-mer, 24-mer, 30-mer, 36-mer, or any number of modules between. Shinoda et al., 2018; Criscuolo et al., 2020. See also US Patent 9,580,714 which is incorporated herein in its entirety.

**[0589]** In some embodiments of the non-guided RNA-binding fusion proteins of the disclosure, the fusion protein comprises at least one RNA-binding protein or RNA-binding portion thereof which is a PUMBY (Pumilio-based assembly) protein. RNA-binding protein PumHD, which has been widely used in native and modified form for targeting RNA, has been engineered into a protein architecture designed to yield a set of four canonical protein modules, each of which targets one RNA base. These modules (*i.e.*, Pumby, for Pumilio-based assembly) are concatenated in chains of varying composition and length, to bind desired target RNAs. In essence, PUMBY is a more simple and modular form of PumHD, in which a single protein unit of PumHD is concatenated into arrays of arbitrary size and binding sequence specificity. The specificity of such Pumby-RNA interactions is high, with undetectable binding of a Pumby chain to RNA sequences that bear three or more

mismatches from the target sequence. *Katarzyna et al., PNAS, 2016; 113(19): E2579-E2588.*

See also US 2016/0238593 which is incorporated by reference herein in its entirety.

**[0590]** In some embodiments of the compositions of the disclosure, the RNA binding protein comprises a Pumilio and FBF (PUF) protein. In some embodiments, the RNA binding protein comprises a Pumilio-based assembly (PUMBY) protein.

**[0591]** In some embodiments of the compositions of the disclosure, at least one of the RNA-binding proteins or RNA-binding portions thereof is a PPR protein. PPR proteins (proteins with pentatricopeptide repeat (PPR) motifs derived from plants) are nuclear-encoded and exclusively controlled at the RNA level organelles (chloroplasts and mitochondria), cutting, translation, splicing, RNA editing, genes specifically acting on RNA stability. PPR proteins are typically a motif of 35 amino acids and have a structure in which a PPR motif is about 10 contiguous amino acids. The combination of PPR motifs can be used for sequence-selective binding to RNA. PPR proteins are often comprised of PPR motifs of about 10 repeat domains. PPR domains or RNA-binding domains may be configured to be catalytically inactive. WO 2013/058404 incorporated herein by reference in its entirety.

**[0592]** In some embodiments, the fusion protein disclosed herein comprises a linker between the at least two RNA-binding polypeptides. In some embodiments, the linker is a peptide linker. In one embodiment, the linker is VDTANGS (SEQ ID NO: 341). In some embodiments, the peptide linker comprises one or more repeats of the tri-peptide GGS. In other embodiments, the linker is a non-peptide linker. In some embodiments, the non-peptide linker comprises polyethylene glycol (PEG), polypropylene glycol (PPG), co-poly(ethylene/propylene) glycol, polyoxyethylene (POE), polyurethane, polyphosphazene, polysaccharides, dextran, polyvinyl alcohol, polyvinylpyrrolidones, polyvinyl ethyl ether, polyacryl amide, polyacrylate, polycyanoacrylates, lipid polymers, chitins, hyaluronic acid, heparin, or an alkyl linker.

**[0593]** In some embodiments, the at least one RNA-binding protein does not require multimerization for RNA-binding activity. In some embodiments, the at least one RNA-binding protein is not a monomer of a multimer complex. In some embodiments, a multimer protein complex does not comprise the RNA binding protein. In some embodiments, the at least one of RNA-binding protein selectively binds to a target sequence within the RNA molecule. In some embodiments, the at least one RNA-binding protein does not comprise an affinity for a second sequence within the RNA molecule. In some embodiments, the at least one RNA-binding protein does not comprise a high affinity for or selectively bind a second

sequence within the RNA molecule. In some embodiments, the at least one RNA-binding protein comprises between 2 and 1300 amino acids, inclusive of the endpoints.

**[0594]** In some embodiments, the at least one RNA-binding protein of the proteins disclosed herein further comprises a sequence encoding a nuclear localization signal (NLS). In some embodiments, a nuclear localization signal (NLS) is positioned at the N-terminus of the RNA binding protein. In some embodiments, the at least one RNA-binding protein comprises an NLS at a C-terminus of the protein. In some embodiments, the at least one RNA-binding protein further comprises a first sequence encoding a first NLS and a second sequence encoding a second NLS. In some embodiments, the first NLS or the second NLS is positioned at the N-terminus of the RNA-binding protein. In some embodiments, the at least one RNA-binding protein comprises the first NLS or the second NLS at a C-terminus of the protein. In some embodiments, the at least one RNA-binding protein further comprises an NES (nuclear export signal) or other peptide tag or secretory signal. In one embodiment, the tag is a FLAG tag.

**[0595]** In some embodiments, the second RNA-binding polypeptide is operably configured to the first RNA-binding polypeptide at the C-terminus of the first RNA-binding polypeptide. In some embodiments, the second RNA-binding polypeptide is operably configured to the first RNA-binding polypeptide at the N-terminus of the first RNA-binding polypeptide

**[0596]** In some embodiments, nucleic acid sequences encoding PUF or PUMBY proteins of the disclosure are codon optimized nucleic acid sequences. In some embodiments, the codon optimized sequence encoding a PUF or PUMBY protein exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased expression in a human subject relative to a wild-type or non-codon optimized nucleic acid sequence.

**[0597]** In some embodiments, nucleic acid sequences encoding PUF or PUMBY proteins of the disclosure are codon optimized nucleic acid sequences. In some embodiments, the codon optimized sequence encoding a PUF or PUMBY protein exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased translation in a human subject relative to a wild-type or non-codon optimized nucleic acid sequence.

**[0598]** In some aspects, a codon optimized nucleic acid sequence encoding a PUF or PUMBY protein exhibits increased stability. In some aspects, a codon optimized nucleic acid sequence encoding a PUF or PUMBY protein exhibits increased stability through increased resistance to hydrolysis. In some embodiments, the codon optimized sequence encoding a

PUF or PUMBY protein exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased stability relative to a wild-type or non-codon optimized nucleic acid sequence. In some embodiments, the codon optimized sequence encoding a PUF or PUMBY protein exhibits at least 5%, at least 10%, at least 20%, at least 30%, at least 50%, at least 75%, at least 100%, at least 200%, at least 300%, at least 500%, or at least 1000% increased resistance to hydrolysis in a human subject relative to a wild-type or non-codon optimized nucleic acid sequence.

**[0599]** In some aspects, a codon optimized nucleic acid sequence encoding a PUF or PUMBY protein can comprise no donor splice sites. In some aspects, a codon optimized nucleic acid sequence encoding a PUF or PUMBY protein can comprise no more than about one, or about two, or about three, or about four, or about five, or about six, or about seven, or about eight, or about nine, or about ten donor splice sites. In some aspects, a codon optimized nucleic acid sequence encoding a PUF or PUMBY protein comprises at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten fewer donor splice sites as compared to a non-codon optimized nucleic acid sequence encoding the PUF or PUMBY protein.

**[0600]** Without wishing to be bound by theory, the removal of donor splice sites in the codon optimized nucleic acid sequence can unexpectedly and unpredictably increase expression of the PUF or PUMBY protein *in vivo*, as cryptic splicing is prevented. Moreover, cryptic splicing may vary between different subjects, meaning that the expression level of the PUF or PUMBY protein comprising donor splice sites may unpredictably vary between different subjects. Such unpredictability is unacceptable in the context of human therapy. Accordingly, the codon optimized nucleic acid sequences which lacks donor splice sites, unexpectedly and surprisingly allows for increased expression of the PUF or PUMBY protein in human subjects and regularizes expression of the PUF or PUMBY protein across different human subjects.

**[0601]** In some aspects, a codon optimized nucleic acid sequence encoding a PUF or PUMBY protein can have a GC content that differs from the GC content of the non-codon optimized nucleic acid sequence encoding the PUF or PUMBY protein. In some aspects, the GC content of a codon optimized nucleic acid sequence encoding a PUF protein is more evenly distributed across the entire nucleic acid sequence, as compared to the non-codon optimized nucleic acid sequence encoding the PUF or PUMBY protein.

**[0602]** Without wishing to be bound by theory, by more evenly distributing the GC content across the entire nucleic acid sequence, the codon optimized nucleic acid sequence exhibits a more uniform melting temperature (“T<sub>m</sub>”) across the length of the transcript. The uniformity of melting temperature results unexpectedly in increased expression of the codon optimized nucleic acid in a human subject, as transcription and/or translation of the nucleic acid sequence occurs with less stalling of the polymerase and/or ribosome.

**[0603]** In some aspects, a codon optimized nucleic acid sequence encoding a PUF or PUMBY protein can have fewer repressive microRNA target binding sites as compared to the non-codon optimized nucleic acid sequence encoding the PUF or PUMBY protein. In some aspects, a codon optimized nucleic acid sequence encoding a PUF or PUMBY protein can have at least one, or at least two, or at least three, or at least four, or at least five, or at least six, or at least seven, or at least eight, or at least nine, or at least ten, or at least ten fewer repressive microRNA target binding sites as compared to the non-codon optimized nucleic acid sequence the PUF protein.

**[0604]** Without wishing to be bound by theory, by having fewer repressive microRNA target binding sites, the codon optimized nucleic acid sequence encoding a PUF or PUMBY protein unexpectedly exhibits increased expression in a human subject.

**[0605]** In some embodiments, PUMBY proteins of the disclosure can target human MAPT. In some embodiments, PUMBY proteins of the disclosure can target and bind exon 10 of a human MAPT. In some embodiments, PUMBY proteins of the disclosure bind UACCAAAGGUGCAG (SEQ ID NO: 390) of exon 10 of a MAPT pre-mRNA. In some embodiments, a MAPT exon 10 targeting PUMBY, PUMBY HO, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLDFRNNRYPNLQLREIAHTEQLVQDQYGSY  
VIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIR  
GHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLE  
HGRPEDKSKIVAEIRGHTEQLVQDQYGNVYVIQHVLEHGRPEDKSKIVAEIRGHTEQL  
VQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLEHGRPED  
KSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQY  
CYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVA  
EIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRH  
VLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTE  
QLVQDQYGNVYVIQHVLEHGRPEDKSKIVAEIRGPHIATLRKYTYGKHILAKLEKYYM  
KNGVDLG (SEQ ID NO: 380). In some embodiments, a MAPT exon 10 targeting PUMBY,

PUMBY HO, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLDFRNNRYPNLQLREIAGHTEQLVQDQYGS  
YVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEI  
RGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVL  
EHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEHGRPEDKSKIVAEIRGHTEQL  
VQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVLEHGRPED  
KSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYG  
CYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVA  
EIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRH  
VLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTE  
QLVQDQYGNVVIQHVLEHGRPEDKSKIVAEIRGHIAHLRKYTYGKHILAKLEKYYM  
KNGVDLG (SEQ ID NO: 381).

[0606] In some embodiments, PUMBY proteins of the disclosure can target human MAPT. In some embodiments, PUMBY proteins of the disclosure can target and bind exon 10 of a human MAPT. In some embodiments, PUMBY proteins of the disclosure bind AUAAGAAGCUGGAU (SEQ ID NO: 391) of exon 10 of a MAPT pre-mRNA. In some embodiments, a MAPT exon 10 targeting PUMBY, PUMBY HQ, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLDFRNNRYPNLQLREIAHTEQLVQDQYGNV  
VIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIR  
GHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLE  
HGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEHGRPEDKSKIVAEIRGHTEQL  
VQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLEHGRPED  
KSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYG  
CYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVA  
EIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQH  
VLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEHGRPEDKSKIVAEIRGHTE  
QLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGPHIALRKYTYGKHILAKLEKYYM  
KNGVDLG (SEQ ID NO: 382). In some embodiments, a MAPT exon 10 targeting PUMBY, PUMBY HQ, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLDFRNNRYPNLQLREIAGHTEQLVQDQYGN  
YVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEI  
RGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVL  
EHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEHGRPEDKSKIVAEIRGHTEQL

VQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVLEHGRPED  
KSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYG  
CYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAE  
IRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHV  
LEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEHGRPEDKSKIVAEIRGHTEQ  
LVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHIATLRKYTYGKHILAKLEKYYMKN  
GVDLG (SEQ ID NO: 383).

[0607] In some embodiments, PUMBY proteins of the disclosure can target human MAPT. In some embodiments, PUMBY proteins of the disclosure can target and bind exon 10 of a human MAPT. In some embodiments, PUMBY proteins of the disclosure bind CUGGAUCUUAGCAA (SEQ ID NO: 392) of exon 10 of a MAPT pre-mRNA. In some embodiments, a MAPT exon 10 targeting PUMBY, PUMBY HR, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLEDFRNNRYPNLQLREIAHTEQLVQDQYGCY  
VIQHVLEHGRPEDKSKIVAEIRGHEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIR  
GHEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHEQLVQDQYGSYVIGHVLE  
HGRPEDKSKIVAEIRGHEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHEQL  
VQDQYGNYSYVIQHVLEHGRPEDKSKIVAEIRGHEQLVQDQYGNYSYVIQHVLEHGRPE  
DKSKIVAEIRGHEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHEQLVQDQY  
GNYSYVIQHVLEHGRPEDKSKIVAEIRGHEQLVQDQYGCYVIQHVLEHGRPEDKSKIV  
AEIRGHEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHEQLVQDQYGSYVIG  
HVLEHGRPEDKSKIVAEIRGHEQLVQDQYGNYSYVIQHVLEHGRPEDKSKIVAEIRGH  
TEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGPHIATLRKYTYGKHILAKLEKYY  
MKNGVDLG (SEQ ID NO: 384). In some embodiments, a MAPT exon 10 targeting  
PUMBY, PUMBY HR, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLEDFRNNRYPNLQLREIAGHTEQLVQDQYGC  
YVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEI  
RGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVL  
EHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQL  
VQDQYGNVYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVYVIQHVLEHGRPE  
DKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQY  
GNVYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIV  
AEIRGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIE  
HVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVYVIQHVLEHGRPEDKSKIVAEIRGH

TEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHIATLRKYTYGKHILAKLEKYY  
MKNGVDLG (SEQ ID NO: 385).

**[0608]** In some embodiments, PUMBY proteins of the disclosure can target human MAPT. In some embodiments, PUMBY proteins of the disclosure can target and bind exon 10 of a human MAPT. In some embodiments, PUMBY proteins of the disclosure bind CGUCCCGGGAGGCG (SEQ ID NO: 393) of exon 10 of a MAPT pre-mRNA. In some embodiments, a MAPT exon 10 targeting PUMBY, PUMBY HS, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLEDFRNNRYPNLQLREIAHTEQLVQDQYGSY  
VIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIR  
GHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLE  
HGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQL  
VQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLEHGRPED  
KSKIVAEIRGHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYG  
SYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAE  
IRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHV  
LEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQ  
LVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGPHIATLRKYTYGKHILAKLEKYYMK  
NGVDLG (SEQ ID NO: 386). In some embodiments, a MAPT exon 10 targeting PUMBY,  
PUMBY HS, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLEDFRNNRYPNLQLREIAGHTEQLVQDQYGS  
YVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEI  
RGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVL  
EHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQL  
VQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVLEHGRPED  
KSKIVAEIRGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYG  
SYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAE  
IRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHV  
LEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQ  
LVQDQYGSYVIRHVLEHGRPEDKSKIVAEIRGHIATLRKYTYGKHILAKLEKYYMKN  
GVDLG (SEQ ID NO: 387).

**[0609]** In some embodiments, PUMBY proteins of the disclosure can target human MAPT. In some embodiments, PUMBY proteins of the disclosure can target and bind exon 10 of a human MAPT. In some embodiments, PUMBY proteins of the disclosure bind

GAGGCGGCAGUGUG (SEQ ID NO: 394) of exon 10 of a MAPT pre-mRNA. In some embodiments, a MAPT exon 10 targeting PUMBY, PUMBY HT, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLDFRNNRYPNLQLREIAHTEQLVQDQYGSY  
VIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEHGRPEDKSKIVAEIR  
GHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEH  
EHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQL  
VQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPED  
KSKIVAEIRGHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSY  
VIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIR  
GHTEQLVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIGHV  
LEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQ  
LVQDQYGSYVIGHVLEHGRPEDKSKIVAEIRGPHIATLRKYTYGKHILAKLEKYYMK  
NGVDLG (SEQ ID NO: 388). In some embodiments, a MAPT exon 10 targeting PUMBY,  
PUMBY HT, comprises the amino acid sequence:

MPKKKRKVEDPKKKRKVDVGRSRLLDFRNNRYPNLQLREIAGHTEQLVQDQYGS  
YVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEHGRPEDKSKIVAEI  
RGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGNVVIQHVLEH  
EHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQL  
VQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPED  
KSKIVAEIRGHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSY  
VIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIRHVLEHGRPEDKSKIVAEIR  
GHTEQLVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHTEQLVQDQYGSYVIEHV  
LEHGRPEDKSKIVAEIRGHTEQLVQDQYGCYVIQHVLEHGRPEDKSKIVAEIRGHTEQ  
LVQDQYGSYVIEHVLEHGRPEDKSKIVAEIRGHIATLRKYTYGKHILAKLEKYYMKN  
GVDLG (SEQ ID NO: 389).

### *Nucleic Acids*

**[0610]** Provided herein are the nucleic acid sequences encoding the gene therapy compositions or RNA-targeting systems for use in gene transfer and expression techniques described herein, including modulating RNA splicing. It should be understood, although not always explicitly stated that the sequences provided herein can be used to provide the expression product as well as substantially identical sequences that produce a protein that has the same biological properties. These “biologically equivalent” or “biologically active” or

“equivalent” polypeptides are encoded by equivalent polynucleotides as described herein. They may possess at least 60%, or alternatively, at least 65%, or alternatively, at least 70%, or alternatively, at least 75%, or alternatively, at least 80%, or alternatively at least 85%, or alternatively at least 90%, or alternatively at least 95% or alternatively at least 98%, identical primary amino acid sequence to the reference polypeptide when compared using sequence identity methods run under default conditions. Specific polypeptide sequences are provided as examples of particular embodiments. Modifications to the sequences to amino acids with alternate amino acids that have similar charge. Additionally, an equivalent polynucleotide is one that hybridizes under stringent conditions to the reference polynucleotide or its complement or in reference to a polypeptide, a polypeptide encoded by a polynucleotide that hybridizes to the reference encoding polynucleotide under stringent conditions or its complementary strand. Alternatively, an equivalent polypeptide or protein is one that is expressed from an equivalent polynucleotide.

**[0611]** The nucleic acid sequences (e.g., polynucleotide sequences) disclosed herein may be codon-optimized which is a technique well known in the art. In some embodiments disclosed herein, exemplary Cas sequences, such as e.g., a nucleic acid sequence encoding SEQ ID NO: 92 (Cas13d known as CasRx) or the nucleic acid sequence encoding SEQ ID NO: 298 (Cas13d known as CasRx), are codon optimized for expression in human cells. Codon optimization refers to the fact that different cells differ in their usage of particular codons. This codon bias corresponds to a bias in the relative abundance of particular tRNAs in the cell type. By altering the codons in the sequence to match with the relative abundance of corresponding tRNAs, it is possible to increase expression. It is also possible to decrease expression by deliberately choosing codons for which the corresponding tRNAs are known to be rare in a particular cell type. Codon usage tables are known in the art for mammalian cells, as well as for a variety of other organisms. Based on the genetic code, nucleic acid sequences coding for, e.g., a Cas protein, can be generated. In some embodiments, such a sequence is optimized for expression in a host or target cell, such as a host cell used to express the Cas protein or a cell in which the disclosed methods are practiced (such as in a mammalian cell, e.g., a human cell). Codon preferences and codon usage tables for a particular species can be used to engineer isolated nucleic acid molecules encoding a Cas protein (such as one encoding a protein having at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type protein) that takes advantage of the codon usage preferences of that particular species. For example, the Cas proteins disclosed herein

can be designed to have codons that are preferentially used by a particular organism of interest. In one example, a Cas nucleic acid sequence is optimized for expression in human cells, such as one having at least 70%, at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 98%, or at least 99% sequence identity to its corresponding wild-type or originating nucleic acid sequence. In some embodiments, an isolated nucleic acid molecule encoding at least one Cas protein (which can be part of a vector) includes at least one Cas protein coding sequence that is codon optimized for expression in a eukaryotic cell, or at least one Cas protein coding sequence codon optimized for expression in a human cell. In one embodiment, such a codon optimized Cas coding sequence has at least 80%, at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type or originating sequence. In another embodiment, a eukaryotic cell codon optimized nucleic acid sequence encodes a Cas protein having at least 85%, at least 90%, at least 92%, at least 95%, at least 96%, at least 97%, at least 98%, at least 99%, or 100% sequence identity to its corresponding wild-type or originating protein. In another embodiment, a variety of clones containing functionally equivalent nucleic acids may be routinely generated, such as nucleic acids which differ in sequence but which encode the same Cas protein sequence. Silent mutations in the coding sequence result from the degeneracy (i.e., redundancy) of the genetic code, whereby more than one codon can encode the same amino acid residue. Thus, for example, leucine can be encoded by CTT, CTC, CTA, CTG, TTA, or TTG; serine can be encoded by TCT, TCC, TCA, TCG, AGT, or AGC; asparagine can be encoded by AAT or AAC; aspartic acid can be encoded by GAT or GAC; cysteine can be encoded by TGT or TGC; alanine can be encoded by GCT, GCC, GCA, or GCG; glutamine can be encoded by CAA or CAG; tyrosine can be encoded by TAT or TAC; and isoleucine can be encoded by ATT, ATC, or ATA. Tables showing the standard genetic code can be found in various sources (see, for example, Stryer, 1988, Biochemistry, 3rd Edition, W.H. Freeman and Co., NY).

**[0612]** “Hybridization” refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson-Crick base pairing, Hoogsteen binding, or in any other sequence-specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi-stranded complex, a single self-hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of a PCR reaction, or the enzymatic cleavage of a polynucleotide by a ribozyme.

[0613] Examples of stringent hybridization conditions include: incubation temperatures of about 25°C to about 37°C; hybridization buffer concentrations of about 6x SSC to about 10x SSC; formamide concentrations of about 0% to about 25%; and wash solutions from about 4x SSC to about 8x SSC. Examples of moderate hybridization conditions include: incubation temperatures of about 40°C to about 50°C; buffer concentrations of about 9x SSC to about 2x SSC; formamide concentrations of about 30% to about 50%; and wash solutions of about 5x SSC to about 2x SSC. Examples of high stringency conditions include: incubation temperatures of about 55°C to about 68°C; buffer concentrations of about 1x SSC to about 0.1x SSC; formamide concentrations of about 55% to about 75%; and wash solutions of about 1x SSC, 0.1x SSC, or deionized water. In general, hybridization incubation times are from 5 minutes to 24 hours, with 1, 2, or more washing steps, and wash incubation times are about 1, 2, or 15 minutes. SSC is 0.15 M NaCl and 15 mM citrate buffer. It is understood that equivalents of SSC using other buffer systems can be employed.

[0614] “Homology” or “identity” or “similarity” refers to sequence similarity between two peptides or between two nucleic acid molecules. Homology can be determined by comparing a position in each sequence which may be aligned for purposes of comparison. When a position in the compared sequence is occupied by the same base or amino acid, then the molecules are homologous at that position. A degree of homology between sequences is a function of the number of matching or homologous positions shared by the sequences. An “unrelated” or “non-homologous” sequence shares less than 40% identity, or alternatively less than 25% identity, with one of the sequences of the present invention.

### ***Cells***

[0615] In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a prokaryotic cell.

[0616] In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a eukaryotic cell. In some embodiments, the cell is a mammalian cell. In some embodiments, the cell is a bovine, murine, feline, equine, porcine, canine, simian, or human cell. In some embodiments, the cell is a non-human mammalian cell such as a non-human primate cell.

[0617] In some embodiments, a cell of the disclosure is a somatic cell. In some embodiments, a cell of the disclosure is a germline cell. In some embodiments, a germline cell of the disclosure is not a human cell.

**[0618]** In some embodiments of the compositions and methods of the disclosure, a cell of the disclosure is a stem cell. In some embodiments, a cell of the disclosure is an embryonic stem cell. In some embodiments, an embryonic stem cell of the disclosure is not a human cell. In some embodiments, a cell of the disclosure is a multipotent stem cell or a pluripotent stem cell. In some embodiments, a cell of the disclosure is an adult stem cell. In some embodiments, a cell of the disclosure is an induced pluripotent stem cell (iPSC). In some embodiments, a cell of the disclosure is a hematopoietic stem cell (HSC).

**[0619]** In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a neuronal cell. In one embodiment, a cell or cells of a patient treated with compositions disclosed herein include, without limitation, central nervous system (neurons), peripheral nervous system (neurons), peripheral motor neurons, and/or sensory neurons. In one embodiment, a neuronal cell is a glial cell.

**[0620]** In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a muscle cell. In some embodiments, a muscle cell of the disclosure is a myoblast or a myocyte. In some embodiments, a muscle cell of the disclosure is a cardiac muscle cell, skeletal muscle cell or smooth muscle cell. In some embodiments, a muscle cell of the disclosure is a striated cell. In one embodiment, a cell or cells of a patient treated with compositions disclosed herein include, without limitation, skeletal muscle (developing and mature muscle fibers and satellite cells), neuromuscular junction, cardiomyocytes, smooth muscle cells, peripheral nervous system (neurons), peripheral motor neurons, and/or sensory neurons.

**[0621]** In some embodiments of the disclosure, a somatic cell is an ocular cell. An ocular cell includes, without limitation, corneal epithelial cells, keratocytes, retinal pigment epithelial (RPE) cells, lens epithelial cells, iris pigment epithelial cells, conjunctival fibroblasts, non-pigmented ciliary epithelial cells, trabecular meshwork cells, ocular choroid fibroblasts, conjunctival epithelial cells. In some embodiments, an ocular cell is a retinal cell or a corneal cell. In one embodiment, a retinal cell is a photoreceptor cell or a retinal pigment epithelial cell. In another embodiment, a retinal cell is a ganglion cell, an amacrine cell, a bipolar cell, a horizontal cell, a Müller glial cell, a rod cell, or a cone cell.

**[0622]** In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a fibroblast or an epithelial cell. In some embodiments, an epithelial cell of the disclosure forms a squamous cell epithelium, a cuboidal cell epithelium, a columnar cell epithelium, a stratified cell epithelium, a pseudostratified columnar cell epithelium or a transitional cell epithelium. In some embodiments, an epithelial cell of the

disclosure forms a gland including, but not limited to, a pineal gland, a thymus gland, a pituitary gland, a thyroid gland, an adrenal gland, an apocrine gland, a holocrine gland, a merocrine gland, a serous gland, a mucous gland and a sebaceous gland. In some embodiments, an epithelial cell of the disclosure contacts an outer surface of an organ including, but not limited to, a lung, a spleen, a stomach, a pancreas, a bladder, an intestine, a kidney, a gallbladder, a liver, a larynx or a pharynx. In some embodiments, an epithelial cell of the disclosure contacts an outer surface of a blood vessel or a vein.

**[0623]** In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a primary cell.

**[0624]** In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is a cultured cell.

**[0625]** In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is in vivo, in vitro, ex vivo or in situ.

**[0626]** In some embodiments of the compositions and methods of the disclosure, a somatic cell of the disclosure is autologous or allogeneic.

### ***Methods of Use***

**[0627]** The disclosure provides a method of modulating pre-mRNA splicing in a subject. The disclosure further provides a method of modulating pre-mRNA splicing comprising administering an RNA-binding composition comprising a nucleic acid sequence encoding a polypeptide, polynucleotide, or nucleoprotein complex capable of binding a target pre-mRNA sequence. The disclosure further provides a method of modulating pre-mRNA splicing comprising administering at least one RNA-binding composition comprising a nucleic acid sequence encoding a polypeptide, polynucleotide, or nucleoprotein complex capable of binding a target pre-mRNA sequence. In some aspects, the method comprises at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 8, or at least 10 RNA binding compositions, wherein each RNA binding composition targets a different pre-mRNA sequence.

**[0628]** The disclosure further provides a method of modulating pre-mRNA splicing comprising administering a composition comprising a nucleic acid sequence encoding (a) an RNA-binding polypeptide or portion thereof, and (b) one or more cognate nucleic acid guide RNAs, wherein each of the one or more cognate nucleic acid guide RNAs are capable of binding a target pre-mRNA sequence.

**[0629]** The disclosure further provides a method of modulating pre-mRNA splicing comprising administering a composition comprising a nucleic acid sequence encoding a non-guided RNA-binding polypeptide or portion thereof comprising a PUF or PUMBY protein capable of binding a target pre-mRNA sequence.

**[0630]** The disclosure provides a method of modifying the level of expression of a gene of the disclosure under conditions suitable for binding of one or more of the guide RNA or the RNA-binding protein or RNA-binding fusion protein (or a portion thereof) to the pre-mRNA molecule.

**[0631]** In some aspects, the RNA-binding composition binds an intronic or exonic sequence of the pre-mRNA. In some aspects, the intronic or exonic sequence comprises an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), or a polypyrimidine tract (PPY). In some aspects, the 5' splice site is a splice donor site. In some aspects, the 3' splice site is a splice acceptor site.

**[0632]** In some aspects, the RNA-binding composition, when bound to the pre-mRNA sequence prevents inclusion of the bound sequence into an mRNA sequence. When the pre-mRNA is processed into a mature mRNA sequence, exonic sequences bound by RNA-binding compositions of the disclosure can be skipped. A skipped exonic sequence is one that is not incorporated into the mature mRNA.

**[0633]** In some aspects, the RNA-binding composition, when bound to the pre-mRNA sequence prevents inclusion of the bound sequence into an mRNA sequence. In some aspects, this results in a truncation of the mature mRNA sequence. In some aspects, the truncated mature mRNA sequence is either not-expressed or expressed at a diminished level.

**[0634]** In some aspects, the intronic or exonic sequence bound by the RNA-binding composition comprises an exon splicing silencer (ESS) motif or an intron splicing silencer (ISS) motif. In some aspects, when an ESS or ISS is bound by RNA-binding compositions of the disclosure, exon inclusion is promoted leading to incorporation of an exon into a mature mRNA sequence.

**[0635]** In some cases, mutated genes and mRNA sequences of the disclosure yield an expressed protein with no activity or diminished activity. The disclosure provides a method of modifying an activity of a protein encoded by an RNA molecule comprising contacting the composition of the disclosure and the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the RNA-binding protein or the fusion protein (or a portion thereof) to the RNA molecule. In some embodiments, the mature mRNA lacking a skipped

exon of the disclosure, when expressed, produced a protein with increased activity. In some embodiments, the protein activity is increased 1- fold or greater. In another embodiment, the level of activity is increased 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold or 10-fold. In another embodiment, the level of activity is increased 10-fold or greater. In another embodiment, the level of activity is increased between 10-fold and 20-fold. In another embodiment, the level of activity is increased 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, 16-fold, 17-fold, 18-fold, 19-fold, or 20-fold.

**[0636]** In some cases, mutated genes and mRNA sequences of the disclosure lead to no expression, very little expression, or reduced expression of a protein encoded by the mutated gene or mRNA sequence. The disclosure provides a method of modifying the expression of a protein encoded by an RNA molecule comprising contacting the composition of the disclosure and the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the RNA-binding protein or the fusion protein (or a portion thereof) to the RNA molecule. In some embodiments, the mature mRNA lacking a skipped exon of the disclosure, when expressed, leads to increased expression of the protein encoded by the mRNA. In some embodiments, the level of protein expression is increased 1- fold or greater. In another embodiment, the level of protein expression is increased 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold or 10-fold. In another embodiment, the level of protein expression is increased 10-fold or greater. In another embodiment, the level of protein expression is increased between 10-fold and 20-fold. In another embodiment, the level of protein expression is increased 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, 16-fold, 17-fold, 18-fold, 19-fold, or 20-fold.

**[0637]** In some cases, mutated genes and mRNA sequences of the disclosure yield an expressed protein with increased activity or hyperactivity. The disclosure provides a method of modifying an activity of a protein encoded by an RNA molecule comprising contacting the composition of the disclosure and the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the RNA-binding protein or the fusion protein (or a portion thereof) to the RNA molecule. In some embodiments, the mature mRNA lacking a skipped exon of the disclosure, when expressed, produced a protein with decreased activity. In some embodiments, the protein activity is decreased 1- fold or greater. In another embodiment, the level of activity is decreased 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold or 10-fold. In another embodiment, the level of activity is decreased 10-fold or greater. In another embodiment, the level of activity is decreased between 10-fold and 20-fold. In another

embodiment, the level of activity is decreased 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, 16-fold, 17-fold, 18-fold, 19-fold, or 20-fold.

**[0638]** In some cases, mutated genes and mRNA sequences of the disclosure lead to increased expression or overexpression of a protein encoded by the mutated gene or mRNA sequence. The disclosure provides a method of modifying the expression of a protein encoded by an RNA molecule comprising contacting the composition of the disclosure and the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the RNA-binding protein or the fusion protein (or a portion thereof) to the RNA molecule. In some embodiments, the mature mRNA lacking a skipped exon of the disclosure, when expressed, leads to decreased expression of the protein encoded by the mRNA. In some embodiments, the level of protein expression is decreased 1- fold or greater. In another embodiment, the level of protein expression is decreased 2-fold, 3-fold, 4-fold, 5-fold, 6-fold, 7-fold, 8-fold, 9-fold or 10-fold. In another embodiment, the level of protein expression is decreased 10-fold or greater. In another embodiment, the level of protein expression is decreased between 10-fold and 20-fold. In another embodiment, the level of protein expression is decreased 11-fold, 12-fold, 13-fold, 14-fold, 15-fold, 16-fold, 17-fold, 18-fold, 19-fold, or 20-fold.

**[0639]** The disclosure provides a method of modifying level of expression of a gene or an RNA molecule of the disclosure or a protein encoded by the RNA molecule comprising contacting the composition of the disclosure and a cell comprising the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the RNA-binding protein or fusion protein (or a portion thereof) to the RNA molecule. In some embodiments, the cell is in vivo, in vitro, ex vivo or in situ. In some embodiments, the composition of the disclosure comprises a vector comprising a guide RNA of the disclosure and an RNA-binding protein or fusion protein of the disclosure. In some embodiments, the vector is an AAV.

**[0640]** The disclosure provides a method of modifying an activity of a protein encoded by an RNA molecule comprising contacting the composition of the disclosure and a cell comprising the RNA molecule under conditions suitable for binding of one or more of the guide RNA or the RNA-binding protein or fusion protein (or a portion thereof) to the RNA molecule.

**[0641]** In some embodiments, the cell is in vivo, in vitro, ex vivo or in situ. In some embodiments, the composition comprises a vector comprising composition comprising a guide RNA of the disclosure and an RNA-binding fusion protein of the disclosure. In some embodiments, the vector is an AAV.

**[0642]** The disclosure provides a method of treating a disease or disorder comprising administering to a subject a therapeutically effective amount of a composition of the disclosure. In one embodiment, the disclosure provides a method of treating monogenic diseases or disorders. Monogenic diseases or disorders are diseases caused by variation in a single gene. In some cases, the variation is at least one mutation to the gene. In some cases, the monogenic disease or disorder arises from more than one mutation in the gene.

**[0643]** Methods of the disclosure can be configured to modulate splicing of any gene in the genome of a subject. In some cases, the gene is a protein-expressing gene. In some embodiments, the gene is USH2A, DMD, SOD1, LRRK2, MAPT, CEP290, SCN1A, SMN1, SMN2, and HTT.

**[0644]** In some embodiments, the monogenic disease or disorder is associated with USH2A. In some embodiments, the USH2A-associated monogenic disease or disorder is Usher Syndrome type II or Retinitis Pigmentosa 39.

**[0645]** In some embodiments, the monogenic disease or disorder is associated with DMD. In some embodiments, the DMD-associated monogenic disease or disorder is Duchenne Muscular Dystrophy or Becker Muscular Dystrophy.

**[0646]** Duchenne's muscular dystrophy is caused by mutations in the DMD gene. DMD is the largest gene in the human genome having 79 exons separated by introns that are up to 250kb. DMD encodes dystrophin protein, which stabilized the plasma membrane in striated muscle. Exon skipping to restore frame (caused by exon 50 deletion frameshift) reduces the severity of DMD to that exhibited by Becker Muscular Dystrophy patients. See FIG. 20.

**[0647]** Duchenne Muscular Dystrophy is caused by large frame-shifting deletions in the DMD gene, leading to insufficient levels of dystrophin protein. Common disease-causing mutations include whole exon 50 or exon 52 deletions. In both cases, skipping of exon 51 would restore reading frame of DMD leading to an internally truncated, but functional dystrophin protein. dCas13d with mutated HEPN domains (to prevent endonucleolytic cleavage) can be programmed to bind exon splicing enhancers (ESEs) and/or other splicing regulatory domains in DMD pre-mRNAs to prevent inclusion of exon 51 during normal pre-mRNA splicing. In addition, Cas13d technology permits expression of multiple guides expressed as a tandem array, which may increase the efficiency and/or specificity of single exon 51 skipping and enable multi-exon skipping, which could be applied to ~65% of DMD patients. See FIG. 21.

[0648] In some embodiments monogenic disease or disorder is associated with SOD1. In some embodiments, the SOD1-associated monogenic disease or disorder is amyotrophic lateral sclerosis (ALS).

[0649] In some embodiments monogenic disease or disorder is associated with LRRK2. In some embodiments, the LRRK2-associated monogenic disease or disorder is Parkinson's disease or Crohn's disease.

[0650] In some embodiments monogenic disease or disorder is associated with MAPT. In some embodiments, the MAPT-associated monogenic disease or disorder is Alzheimer's disease.

[0651] In some embodiments, the monogenic disease or disorder is associated with SMN1 or SMN2. In some embodiments, the SMN2 or SMN1-associated monogenic disease or disorder is spinal muscular atrophy (SMA).

[0652] In some embodiments of the methods of the disclosure, a subject of the disclosure is female. In some embodiments of the methods of the disclosure, a subject of the disclosure is male. In some embodiments, a subject of the disclosure has two XX or XY chromosomes. In some embodiments, a subject of the disclosure has two XX or XY chromosomes and a third chromosome, either an X or a Y.

[0653] In some embodiments of the methods of the disclosure, a subject of the disclosure is a neonate, an infant, a child, an adult, a senior adult, or an elderly adult. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30 or 31 days old. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11 or 12 months old. In some embodiments of the methods of the disclosure, a subject of the disclosure is at least 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100 or any number of years or partial years in between of age.

[0654] In some embodiments of the methods of the disclosure, a subject of the disclosure is a mammal. In some embodiments, a subject of the disclosure is a non-human mammal.

[0655] In some embodiments of the methods of the disclosure, a subject of the disclosure is a human.

[0656] In some embodiments of the methods of the disclosure, a therapeutically effective amount comprises a single dose of a composition of the disclosure. In some embodiments, a therapeutically effective amount comprises a therapeutically effective amount comprises at least one dose of a composition of the disclosure. In some embodiments, a therapeutically

effective amount comprises a therapeutically effective amount comprises one or more dose(s) of a composition of the disclosure.

**[0657]** In some embodiments of the methods of the disclosure, a therapeutically effective amount eliminates a sign or symptom of the disease or disorder. In some embodiments, a therapeutically effective amount reduces a severity of a sign or symptom of the disease or disorder.

**[0658]** In some embodiments of the methods of the disclosure, a therapeutically effective amount eliminates the disease or disorder.

**[0659]** In some embodiments of the methods of the disclosure, a therapeutically effective amount prevents an onset of a disease or disorder. In some embodiments, a therapeutically effective amount delays the onset of a disease or disorder. In some embodiments, a therapeutically effective amount reduces the severity of a sign or symptom of the disease or disorder. In some embodiments, a therapeutically effective amount improves a prognosis for the subject.

**[0660]** In some embodiments of the methods of the disclosure, a composition of the disclosure is administered to the subject via intracerebral administration. In some embodiments, the composition of the disclosure is administered to the subject by an intrastriatal route. In some embodiments, the composition of the disclosure is administered to the subject by a stereotaxic injection or an infusion. In some embodiments, the composition is administered to the brain. In some embodiments of the methods of the disclosure, a composition of the disclosure is administered to the subject locally.

**[0661]** In some embodiments, the compositions disclosed herein are formulated as pharmaceutical compositions. Briefly, pharmaceutical compositions for use as disclosed herein may comprise a protein(s) or a polynucleotide encoding the protein(s), optionally comprised in an AAV, which is optionally also immune orthogonal, in combination with one or more pharmaceutically or physiologically acceptable carriers, diluents or excipients. Such compositions may comprise buffers such as neutral buffered saline, phosphate buffered saline and the like; carbohydrates such as glucose, mannose, sucrose or dextrans, mannitol; proteins; polypeptides or amino acids such as glycine; antioxidants; chelating agents such as EDTA or glutathione; adjuvants (e.g., aluminum hydroxide); and preservatives. Compositions of the disclosure may be formulated for routes of administration, such as e.g., oral, enteral, topical, transdermal, intranasal, and/or inhalation; and for routes of administration via injection or infusion such as, e.g., intravenous, intramuscular, subpial, intrathecal, intraparenchymal, intrathecal, intrastriatal, subcutaneous, intradermal,

intraperitoneal, intratumoral, intravenous, intraocular, and/or parenteral administration. In certain embodiments, the compositions of the present disclosure are formulated for intracerebral or intrastriatal administration.

## EXAMPLES

### Example 1: Screening of inactivated Cas13d nuclease variants

**[0662]** Mutations to Cas13d variants were introduced to deactivate their nuclease activity, thereby producing deactivated Cas13d (dCas13d). Table 48 lists the Cas13d variants that were assessed. FIG. 5A depicts the activity of each dCas13d as measured by their relative ability to cleave Cas13d. FIG. 5B depicts the nuclease activity of active Cas13d variants that do not bear the deactivating mutations.

**[0663]** HEK293 cells were seeded at 20K cells/well in 96-well plates. 100 ng of Cas13d variant, 100 ng of guide RNA and 10 ng of pLenti-GFP were transfected. Each condition was plated in 4 replicates. A plate reader was used to measure GFP at a wavelength of 479 nm. Values were background subtracted (the average of 4 media-only wells) before being normalized to non-targeting (NT). Error bars indicate standard deviation.

Table 48: Deactivated Cas13d variants

| dCas13d | Mutation(s)                    | SEQ ID NO:       |
|---------|--------------------------------|------------------|
| 42      | H841A                          | (SEQ ID NO: 397) |
| 179     | H840A                          | (SEQ ID NO: 399) |
| 189     | H892A                          | (SEQ ID NO: 401) |
| 198     | H841A                          | (SEQ ID NO: 400) |
| 212     | H919A                          | (SEQ ID NO: 396) |
| 208     | H901A                          | (SEQ ID NO: 398) |
| dCasRX  | R238A, H243A, R857A, and H862A | (SEQ ID NO: 402) |

### Example 2: Exon 13 skipping of human USH2A

**[0664]** Patients with USH2A syndrome experience deafness and retinal degeneration. A single base-pair deletion, c.2299 delG, in exon 13 is the most common mutation in USH2A gene, resulting in a truncated and non-functional Usherin protein. There are an estimated 2,000 - 4,000 patients in the United States with the c.2299 delG mutation and approximately 16,000 patients with Exon 13 mutations. Usherin localizes to connecting cilia of mammalian photoreceptors and its absence leads to a ciliary defect and cell death.

**[0665]** Mutations in USH2A are responsible for 10-15% of autosomal recessive retinitis pigmentosa (ARRP). Exon 13 in particular is a hotspot of USH2A mutations. In people having exon 13 USH2A mutations, skipping exon 13 during RNA splicing and processing offers a means to restore healthy USH2A levels and function.

**[0666]** CRISPR Cas RNA guides targeting intronic sequences flanking exon 13 and exon 13 of USH2A pre-mRNA were designed and detailed in Table 49 and depicted in FIG. 3A.

**Table 49: Human USH2A Exon 13 Guide RNA spacer sequences**

| Guide | Target Site                            | Spacer Sequence                             | Exon 13 skipping activity |
|-------|----------------------------------------|---------------------------------------------|---------------------------|
| g1    | Polypyrimidine (PPY) tract (YYY)       | TTTAAAAGGTGAGGATGGGAAAATGA (SEQ ID NO: 403) | partial                   |
| g2    | 3' splice site                         | CCCTAAAGATAAAATATATTTAAAAG (SEQ ID NO: 404) | no                        |
| g3    | 3' splice site                         | CACACCTAAGCCCTAAAGATAAAATA (SEQ ID NO: 405) | no                        |
| g4    | 5' splice site                         | ACATTTCTTTCTTACCTGGTTGACAC (SEQ ID NO: 406) | partial                   |
| g5    | 5' splice site                         | AATGTAATACATTTCTTTCTTACCTG (SEQ ID NO: 407) | no                        |
| g6    | exon splicing enhancer (ESE) motif     | CTGAGCCATGGAGGTTACACTGGCAG (SEQ ID NO: 408) | partial                   |
| g7    | exon splicing enhancer (ESE) motif     | ACACACAGGCACTGGCCACTGATTGG (SEQ ID NO: 409) | yes                       |
| g8    | intronic splicing enhancer (ISE) motif | GATGGTAATAGAGATGTGACTGTAAA (SEQ ID NO: 410) | partial                   |

|     |                                    |                                                |     |
|-----|------------------------------------|------------------------------------------------|-----|
| g9  | exon splicing enhancer (ESE) motif | AGGCACTGGCCACTGATTGGGTCACA<br>(SEQ ID NO: 411) | yes |
| g10 | exon splicing enhancer (ESE) motif | ACTGGCCACTGATTGGGTCACAAATG<br>(SEQ ID NO: 412) | yes |

**[0667]** Cells were treated with dCasRx in combination with guide RNAs from Table 1 targeting exon 13 of a human USH2A minigene and the level of exon 13 skipping induced by each dCasRx complex was assessed. HEK293T cells were transfected with 500 ng of guide-expressing plasmid and 500 ng of dCasRx-expressing plasmid and 50 ng of human USH2A minigene. 48 hours after transfection, RNA was harvested, cDNA was synthesized, 25 cycles of PCR was performed. 500 ng of RNA was subjected to reverse transcription using qScript cDNA synthesis kit. 1 µl of cDNA was used in a 25 cycle, 25 µl GoTaq PCR reaction with oligos annealing to exons 12 and 14 with annealing at 55C and 1 minute extension. The PCR product was run on a D1000 tape (Agilent).

**[0668]** Differential levels of exon 13 skipping was observed depending upon which region of exon 13 was targeted (FIG. 3B). Guide 7, targeting the ESE motif induced almost complete skipping of exon 13 (FIG. 3B and FIG. 3C). Guide 7 targeting the predicted exon splicing enhancer in exon 13 is most effective promoting exon 13 exclusion. Other guides, including those targeting the PPY tract and another predicted ESE in exon 13 are partially effective.

**[0669]** The fraction excluded, depicted in FIG. 3C and correlating to the amount of exon 13 excluded (i.e. skipped) in the resulting USH2A transcripts, is calculated by dividing the intensity of the peak by the sum of the intensities of from the band representing exon 13 included and the band representing exon 13 excluded.

**[0670]** Dose dependency of delivered AAV vectors encoding USH2A pre-mRNA targeting dCas13d complexes was assessed. FIG. 3D is a bar graph depicting dose-dependent exon 13 skipping upon delivery of USH2A pre-mRNA targeting dCas13d complexes with guides g7, g9 and g10 targeting human USHA2A exon 13, showing dose dependent Exon 13 skipping.

**[0671]** FIG. 3E is a series of tape station images depicting exclusion of exon 13 of human USH2A using vectors expressing guide RNA, g7, g9, g10, or a non-targeting guide RNA. Unitary plasmids were transfected at varying doses and RNA was harvested at 48 hr.

**[0672]**

[0673] FIG. 6B is a gel depicting that AAV vector constructs expressing dSeq212 and guides g7, g9 or g10, targeting the ESE of exon 13, were co-transfected with the human USH2A minigene. Guides g7, g9, and g10 induce robust exon 13 skipping.

**Example 3: Impact of inactivated Cas13 nuclease variants on USH2A exon 13 skipping**

[0674] The impact of single point mutant Cas13d orthologs on exon skipping was evaluated. The cas13d orthologs, dSeq42, dSeq179, dSeq189, dSeq198, dSeq212, and dSeq208, were transfected with Cas13d-relevant guide g7 (from Table 49) and the degree of USH2A exon 13 skipping was measured (FIGs 4A and 4B).

[0675] Semiquantitative PCR to exon 12 and exon 14 of the USH2A minigene shows that all catalytically-inactive Cas13d orthologs tested are effective in exon 13 skipping when co-transfected with a the Cas13-specific guide (with g7 spacer sequence; previously known as g16) as represented by the absence of the exon 13-included band at 860bp and the increase in intensity of the exon 13-excluded band at 208bp when compared to a non-targeting guide (with Konermann NT1 spacer sequence; TCACCAGAAGCGTACCATACTCACGA (SEQ ID NO: 445)).

[0676] 500 ng of RNA was used in a 10  $\mu$ L cDNA reaction. 1  $\mu$ L of cDNA was used in a 25  $\mu$ L GoTaq PCR with exon 12 fwd and exon 14 rev primers, with annealing at 55C, 1 minute extension, and 25 cycles. DNA was run on a D1000 tape.

**Example 4: Exon 12 skipping in murine USH2A**

[0677] Mouse and human variants of USH2A share sequence homology. FIG. 6A is an alignment of human USH2A exon 13 and mouse USH2A exon 12. Accordingly, the approach to skipping of exon 12 of USH2A in mice is analogous to the approach to skipping human USH2A exon 13.

[0678] Guide RNAs, mg1 and mg4-mg9, that target different sites of exon 12 of the mouse USH2A as well as intronic sequences are depicted in FIG. 7A.

**Table 50: Murine USH2A Exon 13 Guide RNA spacer sequences**

| Guide       | Spacer Sequence                                 |
|-------------|-------------------------------------------------|
| mg1         | AGACATAGACACTGGCCACTGATGGG (SEQ ID NO: 413)     |
| mg1 – 30 nt | GTAGACATAGACACTGGCCACTGATGGGGT (SEQ ID NO: 414) |
| mg4         | GCTGCAGTCACAGACCTCACAAGCAC (SEQ ID NO: 415)     |
| mg5         | CTGGAAACAAAACACAATGGAAAGGA (SEQ ID NO: 416)     |

|             |                                                 |
|-------------|-------------------------------------------------|
| mg6         | GAAATTGCAGCGATTGCATCTAAGCC (SEQ ID NO: 417)     |
| mg7         | AGCGATTGCATCTAAGCCCTGGAAAC (SEQ ID NO: 418)     |
| mg7 – 30 nt | GCAGCGATTGCATCTAAGCCCTGGAAACAA (SEQ ID NO: 419) |
| mg8         | TCTCTGCAAGAGTCACACTTAAGTCC (SEQ ID NO: 420)     |
| mg9         | AGTCCCAGGCTGGGACCCAGCCTTGC (SEQ ID NO: 421)     |

[0679] FIG. 7B is a tape station image depicting the effect of mouse USH2A pre-mRNA targeting dCasRX complexes on USH2A minigene splicing, wherein the various complexes each bear mg1 and an additional guide: mg4-mg9. Combination of guides mg1 and mg7 targeting exon 12 ESE and the 5' end of exon 12, respectively, induced robust exon 12 skipping.

#### **Example 5: Development of methods for exon skipping in mice**

[0680] As detailed in Example 3 and FIG. 6D, the use of two dCas13 proteins with guides that target distinct regions of an exon yields robust exon skipping. FIG. 8A is a gel image demonstrating the robust skipping achieved by combining guides.

[0681] FIG. 8B is a series of microscope images detecting splice junctions of mouse USH2A minigene in cosm6 cells transfected with 250 ng of guides mg1 and mg7, 500 ng dCasRx and 50 ng mouse minigene in 12 well format. Cells were trypsinized and seeded in chamber slides 24 hr before fixation with PFA. Cells not transfected with minigene show no detection of splice junctions, cells treated with a non-targeting (NT) guide show exon 11/exon12 splice junction, cells treated with both mg1 and mg7 show less exon 11/exon 12 junction and presence of exon 11/exon 13 splice junction (indicative of exon 12 skipping) .

#### **Example 6: Development of AAV vectors for delivering exon skipping dCas13d complexes**

[0682] AAV vectors were developed for delivery of USH2A exon skipping complexes to both mice and human.

[0683] For mice, an exemplary AAV vector construct encoding mg1 and mg7 guide RNAs under the control of a human U6 (hU6) promoter and a dCas13d-NLS (dCas13d tethered to a nuclear localization signal) under the control of an EFS promoter for the targeting of exon 12 of mouse USH2A is depicted in FIG. 9A.

[0684] For human, an exemplary AAV vector construct encoding a g7 guide RNA under the control of a human U6 (hU6) promoter and a dCas13d-NLS (dCas13d tethered to a nuclear localization signal) under the control of a cone/rod specific promoter for the targeting of exon 13 of human USH2A is depicted in FIG. 9B.

**Example 7: Dose response of AAV vectors targeting mouse USH2A exon 12**

[0685] The dose-dependency of mouse USH2A exon 12 skipping was assessed by transfecting cells with AAV vectors at doses of 1 µg, 500 ng, 250 ng, or 100 ng.

[0686] AAV vectors expressing dual guide RNAs, mg1 and mg7, were assessed. Vector A01270 (SEQ ID NO: 598) containing dSeq212 under an EFS promoter and guides mg1 and mg7 each controlled by their own hU6 promoter. Vector A01271 (SEQ ID NO: 509) containing dSeq212 under an EFS-UBB promoter and guides mg1 and mg7 each controlled by their own hU6 promoter. Vector A01285 (SEQ ID NO: 510) containing dSeq212 under an EFS-UBB promoter and guides mg1 and mg7 under the control of a single hU6 promoter. FIG. 10 depicts splicing modulation of exon 12 of mouse USH2A. Unitary plasmids were transfected at varying doses and RNA was harvested at 48 hr. Cells were transfected with 1 µg, 500 ng, 250 ng, 100 ng of AAV vector (total DNA transfected: 1 µg per condition) and 50 ng mouse minigene for all conditions.

**Example 8: Exon skipping in DMD for treating Duchenne Muscular Dystrophies**

[0687] Duchenne muscular dystrophy is caused by mutations in the *DMD* gene. DMD is the largest gene in human genome; 79 exons separated by introns that are up to 250 kb. DMD encodes dystrophin protein, which stabilizes the plasma membrane in striated muscle. A hotspot of mutations is found between exons 43 and 55. Distinct mutations in DMD lead to different types of Duchenne Muscular Dystrophy as detailed in Table 51.

**Table 51: Disease causing mutations in DMD**

|               | <b>Duchenne Muscular Dystrophy</b>                                     | <b>Becker Muscular Dystrophy</b>                                        |
|---------------|------------------------------------------------------------------------|-------------------------------------------------------------------------|
| Mutation      | Deletion of out of frame exons                                         | Deletion of in frame exons                                              |
| Protein Level | Low to no dystrophin                                                   | 10-40% protein                                                          |
| Severity      | Symptomatic by 2-3, can be ambulatory to 12, death by 20 years of life | Later onset, ambulatory to at least 15, survival to 30-40 years of life |

[0688] CRISPR Cas RNA guides targeting exon 51 of DMD pre-mRNA were designed and detailed in Table 52 and depicted in FIG. 11A.

**Table 52: DMD Exon 51 Guide RNA spacer sequences**

| Guide | Spacer Sequence                             |
|-------|---------------------------------------------|
| g1    | ggaaaaaagaagaaaaagaaaaTTA (SEQ ID NO: 422)  |
| g2    | GGAGCTAAAATATTTGGGTTTTGC (SEQ ID NO: 423)   |
| g3    | GTCTGAGTAGGAGCTAAAATATTTG (SEQ ID NO: 424)  |
| g4    | TGAGTAGGAGCTAAAATATTTGGGT (SEQ ID NO: 425)  |
| g5    | AACATCAAGGAAGATGGCATTCTAG (SEQ ID NO: 426)  |
| g6    | ACATCAAGGAAGATGGCATTCT (SEQ ID NO: 427)     |
| g7    | CCTCTGTGATTTTATAACTTGAT (SEQ ID NO: 428)    |
| g8    | TCATTTTTTCTCATACCTTCTGCTTG (SEQ ID NO: 429) |
| g9    | CTTTTATCATTTTTTCTCATACCTTC (SEQ ID NO: 430) |
| g10   | TTTTTCTCATACCTTCTGCTTGATGA (SEQ ID NO: 431) |
| g11   | CTGAGTAGGAGCTAAAATATTTGGG (SEQ ID NO: 432)  |
| g12   | AGTAGGAGCTAAAATATTTGGG (SEQ ID NO: 433)     |
| g13   | CAGAGTAACAGTCTGAGTAGGAGCTA (SEQ ID NO: 434) |
| g14   | GTAACAGTCTGAGTAGGAGCTA (SEQ ID NO: 435)     |

[0689] Exon 51 skipping was evaluated using dCasRx and dSeq189 with guides from table 52. Guide g5 (Table 52) induced robust exon 51 skipping of a minigene comprising exogenous DMD exons 50, 51, and 52 (FIG. 11B).

**Example 9: Multiple guides to target multiple exon skipping in DMD**

[0690] A multi-exon skipping approach to restore dystrophin levels by skipping exons 45-50 or exons 46-50 was assessed (FIG. 12).

[0691] Guides that target exons 45 and 51 of DMD were developed and assessed alone or in combination with g5 (FIG. 13B) for their ability to induce skipping of exons 45-51.

[0692] Guide 45 sequence: TGCCGCTGCCCAATGCCATCCTGGAG (SEQ ID NO: 436).

[0693] Guide 51 sequence: CCTCTGTGATTTTATAACTTGATCAA (SEQ ID NO: 437).

**Example 10: Exon 10 skipping in human MAPT minigene using dCasRx**

[0694] Mutations to MAPT, which is the gene encoding the tau protein, are associated with a variety of tauopathies including but not limited to Alzheimer's disease, progressive supranuclear palsy, corticobasal syndrome, some frontotemporal dementias, and chronic traumatic encephalopathy. Skipping exons bearing tauopathy-associated mutations offers a means to restoring functional levels of tau in subjects.

[0695] CRISPR Cas RNA guides targeting exon 10 of MAPT pre-mRNA were designed and detailed in Table 54.

**Table 54: MAPT Exon 10 Guide RNAs**

| Guide | Spacer Sequence                              |
|-------|----------------------------------------------|
| g1    | aattatctgcacctttgtagccaga (SEQ ID NO: 438)   |
| g2    | taattatctgcacctttgtagccag (SEQ ID NO: 439)   |
| g3    | tgtgaagggtactcacactgccgcctc (SEQ ID NO: 440) |

[0696] MAPT minigene JO, expressing exon 9, 10 and 11 of MAPT separated by truncated introns were used to evaluate exon skipping using dCasRx with guides g1-g3. g3 induced robust exon 10 skipping using MAPT minigene (FIG. 14). MAPT minigene JP was also used to show exon 10 skipping.

**Example 11: Exon 10 skipping in human MAPT minigene using PUMBYs**

[0697] PUMBYs were developed and evaluated for their ability to induce exon 10 skipping in MAPT JO minigene (FIG. 15).

[0698] The impact of targeting exon 10 of a human MAPT minigene with PUMBY proteins was evaluated (FIG. 15). A positive control is dCasRx complexed with guide g3 from FIG. 14. The gel image depicts a 197 bp band corresponding to a transcript comprising exon 10. A second 104 bp band is also depicted corresponding to the absence of exon 10, indicative of exon 10 skipping. All PUMBYs evaluated induced at least partial exon 10 skipping, with PUMBY HR and HS producing the most robust exon 10 skipping.

**Example 12: Splicing mediated knockdown of SOD1 and LRRK2**

[0699] SOD1 mutations are responsible for 20% of familial ALS and 2% of Sporadic ALS. C9orf72 GGGGCC expansions are responsible for 35% of familial ALS and 10% of sporadic ALS. Knocking down SOD1 by approximately 80% increases life span and resolves motor problems in SOD1 G37R mutant mice. Therefore, the induction of exon skipping can lead to

degradation of SOD1 and LRRK2 transcripts. FIG. 16A is a schematic depicting a SOD1 cDNA comprising exons 1-5. FIG. 17 is a schematic depicting an LRRK3 cDNA comprising exons 1-4.

[0700] CRISPR Cas RNA guides targeting exons 3 and 2 of SOD1 pre-mRNA were designed and detailed in Table 55.

**Table 55: SOD1 Exon 2 and Exon 3 Guide RNAs**

| Guide | Spacer Sequence                             |
|-------|---------------------------------------------|
| g1    | TTCATCCTTTGGCCACCGTGTTTC (SEQ ID NO: 441)   |
| g2    | TTCCCCACACCTTCACTGGTCCATTA (SEQ ID NO: 442) |
| g3    | TTCAGTCAGTCCTTTAATGCTTCCCC (SEQ ID NO: 443) |
| g4    | GCTGTATTATCTCCAAACTCATGAAC (SEQ ID NO: 444) |

[0701] The impact of guides g1-g4, targeting exons 2 and 3 of endogenous SOD1, complexed with dCasRx is shown in FIG. 16C. The gels depict a 224 bp band corresponding to a transcript comprising exons 1-4. A second 154 bp band is also depicted corresponding to the absence of exon 3, indicative of exon 3 skipping. A third 127 bp band is also depicted corresponding to the absence of exon 2, indicative of exon 2 skipping. A fourth 57 bp band is also depicted corresponding to the absence of exons 2 and 3, indicative of exon 2 and 3 skipping. Skipping of exons 2 and/or 3 is will introduce a premature termination codon that is predicted to lead to nonsense-mediated mRNA decay (NMD).

**Example 13: AAV8-dCas13d tandem guide vector drives robust Cas13d expression and exon 12 skipping in mouse photoreceptors**

[0702] AAV5, 8, 9 or Rh10 vectors encoding GFP were delivered by subretinal injection to wt BALB/C mice at 1e9 vg/eye (FIG. 18A). Histology was performed 4-weeks post injection using 12 µm retinal cryosections. Native EGFP expression was compared across serotypes to assess spread and photoreceptor targeting. Sub-retinal delivery of AAV serotypes identified AAV8 as the most effective for rod/cone photoreceptor transduction and vector spread (FIG. 18B).

[0703] AAV8-U6-mg1/mg7-dCas13d vector was manufactured in-house and delivered by sub retinal injection to wt BALB/C mice at 1e9 vg/eye (FIG. 18C). Histology was performed

4-weeks post injection. dCas13d mRNA expression was assessed by RNAscope probes specific for Ush2a (FIG. 18D). Ush2a splicing was assessed by duplex BaseScope with probes for spliced (Exon11/13) and intact (Exon11/12) mRNA (FIG. 18D - FIG. 18E).

**Example 14: Sub retinal delivery of AAV8-dCAS13d+mg1/mg7 in mouse retina shows preservation of retinal structure and minimal T-cell infiltration**

[0704] AAV8-U6-mg1/mg7-dCas13d vector was manufactured in-house and delivered by sub-retinal injection to wt BALB/C mice at 1e9 vg/eye (FIG. 19A). Immunohistology was performed 4-weeks post injection on 12  $\mu$ m retinal cryosections. Retina were imaged using immunofluorescence microscopy (FIG. 19B). Cone photoreceptors were immunolabeled with the anti-cone opsin Antibody (Ab). Rod photoreceptors were immunolabeled with the anti-rhodopsin Ab. T Cells were immunolabeled using the anti-CD3 Ab.

**Example 15: Identification and Use of Multi-targeting guides for efficient and specific DMD exon 51 skipping**

[0705] Guides were designed to target exon 51 splicing enhancers, cloned into 30nt Seq212-DR (for guides 15-40) and 36 nt Seq212-DR (for guides 41-42), transfected alongside dSeq212 (4aa mutation) at a 1:1 ratio of guide:dSeq212 into a cell line which expresses DMD from the endogenous locus. (For example, dual guide-targeting experiments, guide concentration was halved such that total guide ratio to dSeq212 remained 1:1, etc.) RNA was extracted, reverse transcribed and subjected to semi-quantitative PCR with primers annealing to exon 49 and exon 53 to detect exon 51 skipping.

*Guide screen identifies targeting regions that result in exon 51 skipping:*

[0706] dSeq212 (4aa mutation) programmed with g5, g17, g38, g42 result in reproducible exon 51 skipping. Multiple exon 51 targeting regions (enriched with predicted SRSF binding sites) can be targeted by dSeq212 to result in efficient exon 51 skipping. Experiment performed in duplicate, exon 49 fwd, exon 53 reverse, 35x cycles, single PCR optimization. See FIGs 22A-22C. *Combining guides that target different ESEs leads to enhanced exon 51 skipping:*

[0707] Multi-targeting guide pair combinations utilizing guide numbers 5, 17, 38, and 42 were selected: g5/g17, g5/g42, g38/g17, g38/g42 and compared to single guide targeting. See FIG. 23A. Lower levels of exon skipping were observed with single guides. Total guide transfected with 500ng (single targeting 500ng; dual targeting 250ng of each guide), single PCR was performed with exon 49 fwd and exon 53 reverse for 35 cycles at an optimal temperature of 58 degrees Celsius. Guide combos g5/g17, g5/g42, g38/g17, g38/g42 which target different ESEs in exon 51 lead to enhanced exon 51 skipping. See FIG. 23B and 23C.

*Guide combos for DMD-targeting systems:*

[0708] Multi-targeting guide pair combinations such as g5/g17, g5/g42, g38/g17, g38/g42 and dSEQ212 (4aa mutation) systems are delivered to achieve higher levels of exon 51 skipping.

**Example 16: Multi-exon skipping (exons 45 through 55) are used to treat 65% of DMD patients**

[0709] Identification of combined guides for targeting additional exons for skipping. See FIG. 24.

**Example 17: Multiple guide targeting of exon 51 of DMD leads to restoration of dystrophin expression**

[0710] Deletion of exon 52 from DMD leads to no dystrophin protein expression. Restoring dystrophin expression can lead to the improvement of symptoms or treatment of Duchenne muscular dystrophy. Without wishing to be bound by theory, skipping exon 51 of the exon 52 deletion DMD gene can restore dystrophin protein expression levels. AAV vector A03278 comprising dCas13d dSeq212 having 4aa mutation is programmed to bind two exon splicing enhancers in exon 51 of human DMD. The 2 guides are expressed from a tandem array, under the hU6 promoter.

[0711] Myotubes having DMD lacking exon 52 were transduced with A03278 for 7 days before harvest and RNA extraction (for cDNA synthesis and semi-quantitative PCR) and cell fixation (for IF).

[0712] FIG. 25A demonstrates that treatment of myotubes with the A03278 vector results in efficient exon 51 skipping. As measured by semi-quantitative PCR, 60% exon 51 skipping is observed relative to a control sample (FIG. 25B).

[0713] DMD lacking exon 52 leads to no dystrophin expression (FIG. 25C). Upon treatment with the vector, thereby inducing exon 51 skipping, dystrophin protein expression is rescued.

**Example 18: dCas13d programmed to target an intronic splicing silencer in intron 7 of SMN2 leads to increased inclusion of exon 7**

[0714] Treatment for SMA includes the restoration of protein expression from SMN2 when SMN1 is dysfunctional. Exon 7 of SMN2 is normally excluded during pre-mRNA splicing due to a SNP in an exon splicing enhancer. Engineering the inclusion of exon 7 will restore expression from SMN2. To test if dCas13d constructs comprising guides targeting to promote exon inclusion we designed guides that would bind and block an intronic splicing silencer in intron 7 of SMN2 (Table B and FIG. 26A) and tested them on a minigene of SMN2 (FIG.

26B). (Semiquantitative PCR was performed and products run on a D1000 tape (FIG. 26C). The bottom band represents exon 7 excluded and the top band represents exon 7 included. The amount of exon 7 inclusion was quantified by PCR (FIG. 26D) and demonstrated that that guides 1, 2 and 3 result in an increase of exon 7 inclusion.

## CLAIMS

We Claim:

1. A pre-mRNA sequence binding composition comprising a nucleic acid sequence encoding (a) an RNA-binding polypeptide or portion thereof, and (b) one or more cognate nucleic acid guide RNA, wherein each of the one or more cognate nucleic acid guide RNA are capable of binding a target pre-mRNA sequence,  
wherein the target pre-mRNA sequence comprises an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), o a polypyrimidine tract (PPY), exon splicing silencer (ESS) motif, or an intron splicing silencer (ISS) motif.
2. A pre-mRNA sequence binding composition comprising a nucleic acid sequence encoding a PUF or PUMBY polypeptide capable of binding a target pre-mRNA sequence;  
wherein the target pre-mRNA sequence comprises an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), o a polypyrimidine tract (PPY), exon splicing silencer (ESS) motif, or an intron splicing silencer (ISS) motif.
3. The composition of claim 1 or 2, wherein the 5' splice site is a splice donor site.
4. The composition of claim 1 or 2, wherein the 3' splice site is a splice acceptor site.
5. The composition of claim 1, wherein the composition comprises at least two cognate nucleic acid guides.
6. The composition of any one of claims 1-5, wherein the polypeptide of is a fusion protein.
7. The composition of claim 5, wherein the fusion protein comprises a peptide capable of modulating splicing.
8. The composition of claim 6, wherein the peptide capable of modulating splicing binds RNA.

9. The composition of claim 7, wherein the peptide capable of modulating splicing is RBFOX1, RMP38, hnRNPA1, or AF2.
10. The composition of claim 1, wherein the RNA-binding polypeptide is a catalytically inactivated Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)-associated (Cas) polypeptide.
11. The composition of claim 10, wherein the Cas polypeptide is a catalytically inactivated or dead Cas13d (dCas13d) polypeptide.
12. The composition of claim 11, wherein the Cas13d polypeptide comprises an amino acid sequence set forth in SEQ ID NO: 396, SEQ ID NO: 397, SEQ ID NO: 398, SEQ ID NO: 399, SEQ ID NO: 400, SEQ ID NO: 401, or SEQ ID NO: 12307.
13. The composition of claim 11, wherein the Cas13d polypeptide comprises an amino acid sequence set forth in SEQ ID NO: 396 or SEQ ID NO: 397.
14. The composition of claim 2, wherein the PUF or PUMBY comprises an amino acid sequence set forth in SEQ ID NO: 380, SEQ ID NO: 381, SEQ ID NO: 382, SEQ ID NO: 383, SEQ ID NO: 384, SEQ ID NO: 385, SEQ ID NO: 386, SEQ ID NO: 387, SEQ ID NO: 388, or SEQ ID NO: 389.
15. The composition of any one of claims 1 or 2, wherein the target pre-mRNA sequence comprises a sequence encoding for USH2A.
16. The composition of claim 15, wherein the sequence encoding for USH2A comprises a mutation.
17. The composition of claim 16, wherein the mutation in USH2A occurs in exon 13.
18. The composition of any one of claims 15-16, wherein the pre-mRNA sequence comprises SEQ ID NO: 467.

19. The composition of any one of claims 15-16, wherein the cognate nucleic acid guide RNA comprises any one of SEQ ID NOs 403-412.
20. The composition of any one of claims 1 or 2, wherein the target pre-mRNA sequence comprises a sequence encoding for DMD.
21. The composition of claim 20, wherein the sequence encoding for DMD comprises a mutation.
22. The composition of claim 24, wherein the mutation in DMD occurs in at least one of exon 42, exon 43, exon 44, exon 45, exon 46, exon 47, exon 48, exon 49, exon 50, exon 51, exon 52, exon 53, exon 54, and exon 55.
23. The composition of any one of claims 20-22, wherein the pre-mRNA sequences comprises SEQ ID NO: 471.
24. The composition of any one of claims 20-22, wherein the cognate nucleic acid guide RNA comprises any one of SEQ ID NOs 422-437, 12272-12276, 12285, 12288-12306, or 12311-12312.
25. The composition of any one of claims 1 or 2, wherein the target pre-mRNA sequence comprises a sequence encoding for MAPT.
26. The composition of claim 25, wherein the sequence encoding for MAPT comprises a mutation.
27. The composition of claim 26, wherein the mutation in MAPT occurs in exon 10.
28. The composition of any one of claims 25-27, wherein the pre-mRNA sequences comprises SEQ ID NO: 479.
29. The composition of any one of claims 25-27, wherein the cognate nucleic acid guide RNA comprises any one of SEQ ID NOs 438-440.

30. The composition of claim 25, wherein a PUMBY protein binds a MAPT exon 10 target sequence set forth in any one of SEQ ID NO: 390, SEQ ID NO: 391, SEQ ID NO: 392, SEQ ID NO: 393, of SEQ ID NO: 394.
31. The composition of any one of claims 1 or 2, wherein the nucleic acid sequence comprises at least one promoter.
32. The composition of claim 31, wherein the at least one promoter is a constitutive promoter or a tissue-specific promoter.
33. The composition of claim 32, wherein the at least one promoter is selected from the group consisting of a tCAG promoter, an EFS promoter, and an EFS-UBB promoter.
34. The composition of claim 31, wherein the nucleic acid sequence comprises two promoters.
35. The composition of claim 34, wherein a first promoter controls expression of the one or more cognate nucleic acid guide RNA and a second promoter controls expression of the RNA-binding polypeptide or portion thereof.
36. A vector comprising the composition of any one of claims 1 or 2.
37. The vector of claim 36, wherein the vector is selected from the group consisting of: adeno-associated virus, retrovirus, lentivirus, adenovirus, nanoparticle, micelle, liposome, lipoplex, polymersome, polyplex, and dendrimer.
38. A cell comprising the vector of claim 36.
39. A method of modulating pre-mRNA splicing comprising administering a composition comprising a nucleic acid sequence encoding (a) an RNA-binding polypeptide or portion thereof, and (b) one or more cognate nucleic acid guides, wherein each of the one or more cognate nucleic acid guides are capable of binding a target pre-mRNA sequence.

40. A method of modulating pre-mRNA splicing comprising administering a composition comprising a nucleic acid sequence encoding a non-guided RNA-binding polypeptide or portion thereof comprising a PUF or PUMBY protein capable of binding a target pre-mRNA sequence.
41. The method of any one of claims 39 or 40, wherein the pre-mRNA sequence comprises an exon splicing enhancer (ESE) motif, an intronic splicing enhancer (ISE) motif, a 5' splice site, a 3' splice site, a branchpoint adenosine (BPA), or a polypyrimidine tract (PPY).
42. The method of claim 41, wherein the 5' splice site is a splice donor site.
43. The method of 41, wherein the 3' splice site is a splice acceptor site.
44. The method of any one of claims 39 or 40, wherein the pre-mRNA sequence comprises an exon splicing silencer (ESS) motif or an intron splicing silencer (ISS) motif.
45. The method of any one of claims 39 or 40, wherein binding the pre-mRNA sequence prevents inclusion of the bound sequence into an mRNA sequence.
46. The method of any one of claims 39 or 40, wherein binding the pre-mRNA sequence promotes inclusion of the bound sequence into an mRNA sequence.
47. The method of any one of claims 39 or 40, wherein binding the pre-mRNA sequence produces a truncated mRNA sequence.
48. The method of any one of claims 39-47, wherein the mRNA sequence is expressed into a protein.
49. The method of claim 48, wherein the protein expressed from the mRNA is expressed at an increased level compared to a wild-type mRNA.
50. The method of claim 48, wherein the protein expressed from the mRNA is expressed at a decreased level compared to a wild-type mRNA.

51. The method of claim 48, wherein the protein expressed from the protein has increased activity compared to a wild-type mRNA.
52. The method of claim 48, wherein the protein expressed from the protein has decreased activity compared to a wild-type mRNA.
53. A method of treating a disease or disorder associated with a gene mutation, wherein the mutation occurs in an exon of a pre-mRNA sequence, comprising administering a composition of any one of claims 1-35.
54. The method of claim 53, wherein the target pre-mRNA sequence comprises a sequence encoding for USH2A.
55. The method of claim 54, wherein the sequence encoding for USH2A comprises a mutation.
56. The method of claim 55, wherein the mutation in USH2A occurs in exon 13.
57. The method of claim 53, wherein the target pre-mRNA sequence comprises a sequence encoding for DMD.
58. The method of claim 57, wherein the sequence encoding for DMD comprises a mutation.
59. The method of claim 58, wherein the mutation in DMD occurs in at least one of exon 42, exon 43, exon 44, exon 45, exon 46, exon 47, exon 48, exon 49, exon 50, exon 51, exon 52, exon 53, exon 54, and exon 55.
60. The method of claim 53, wherein the disease or disorder is Duchenne muscular dystrophy, amyotrophic lateral sclerosis, or Usher syndrome type II.

FIG. 1A

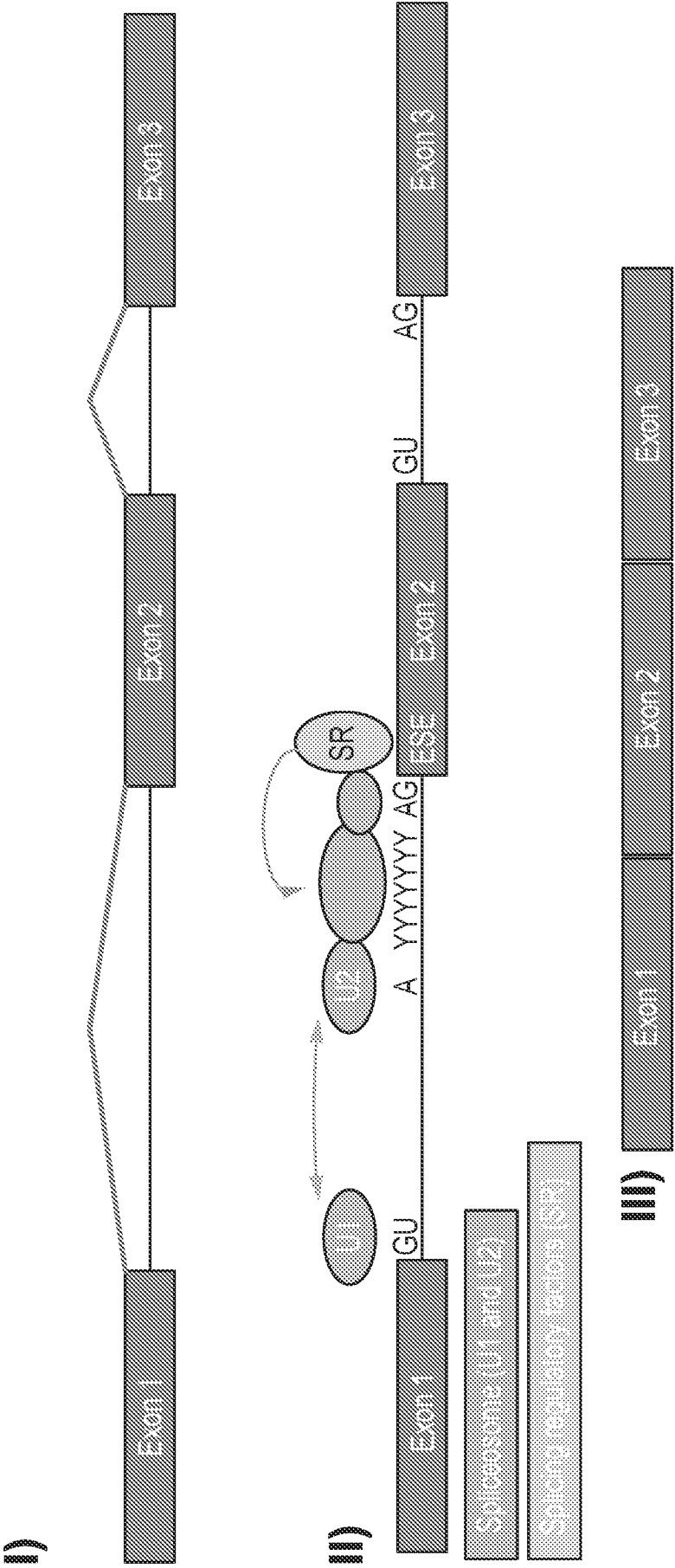


FIG. 1B

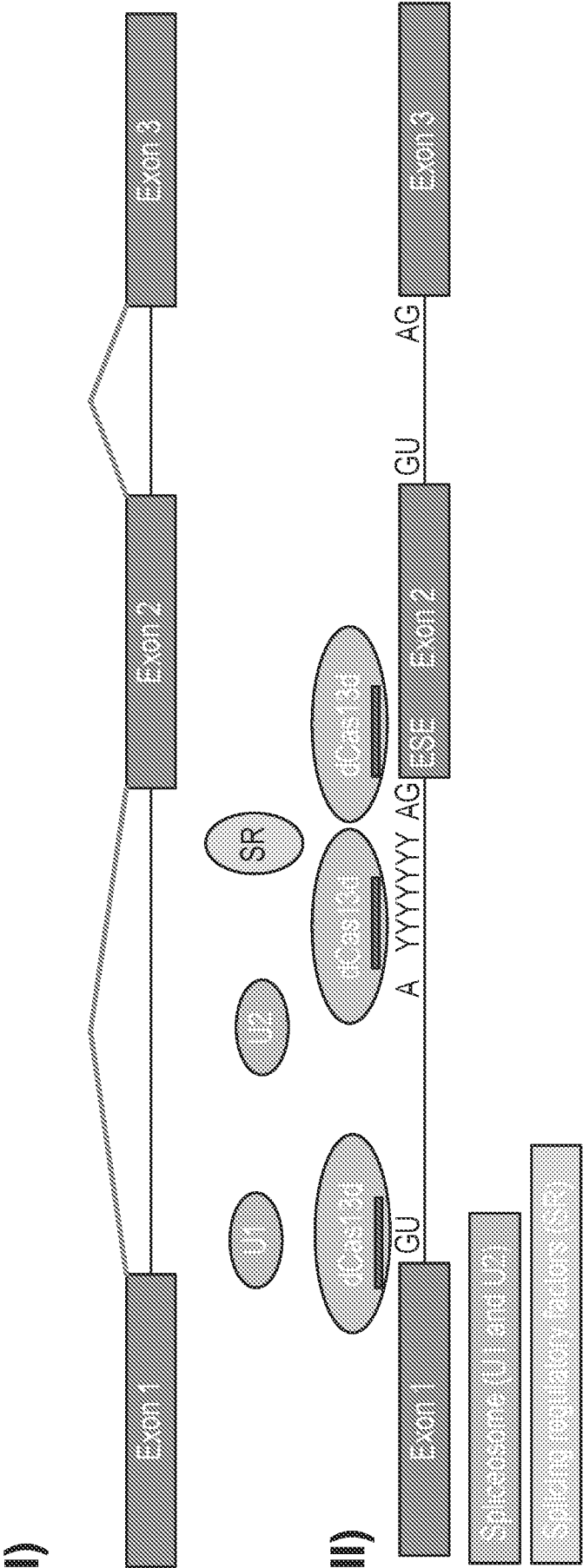


FIG. 2

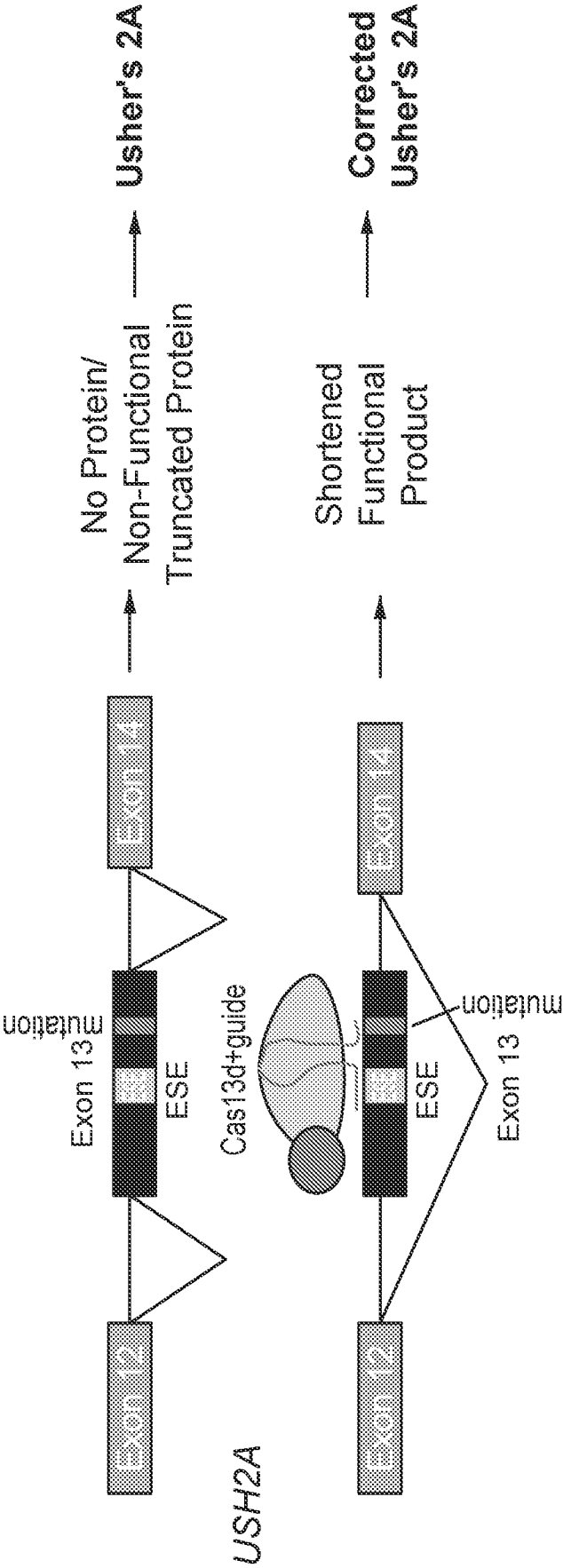


FIG. 3A

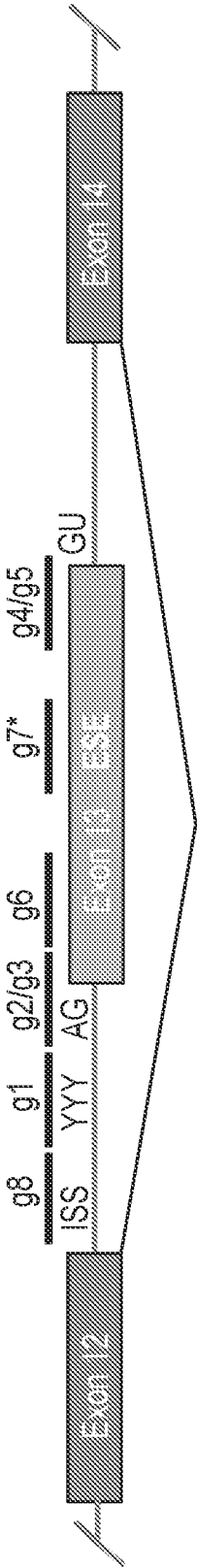


FIG. 3B

dCasRx + 50 ng minigene

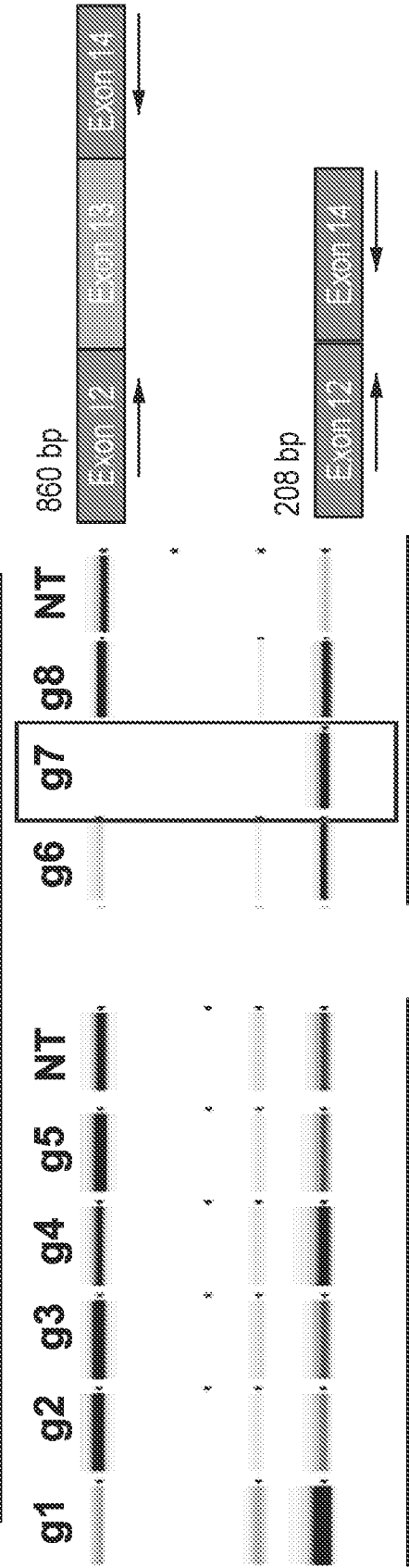


FIG. 3C

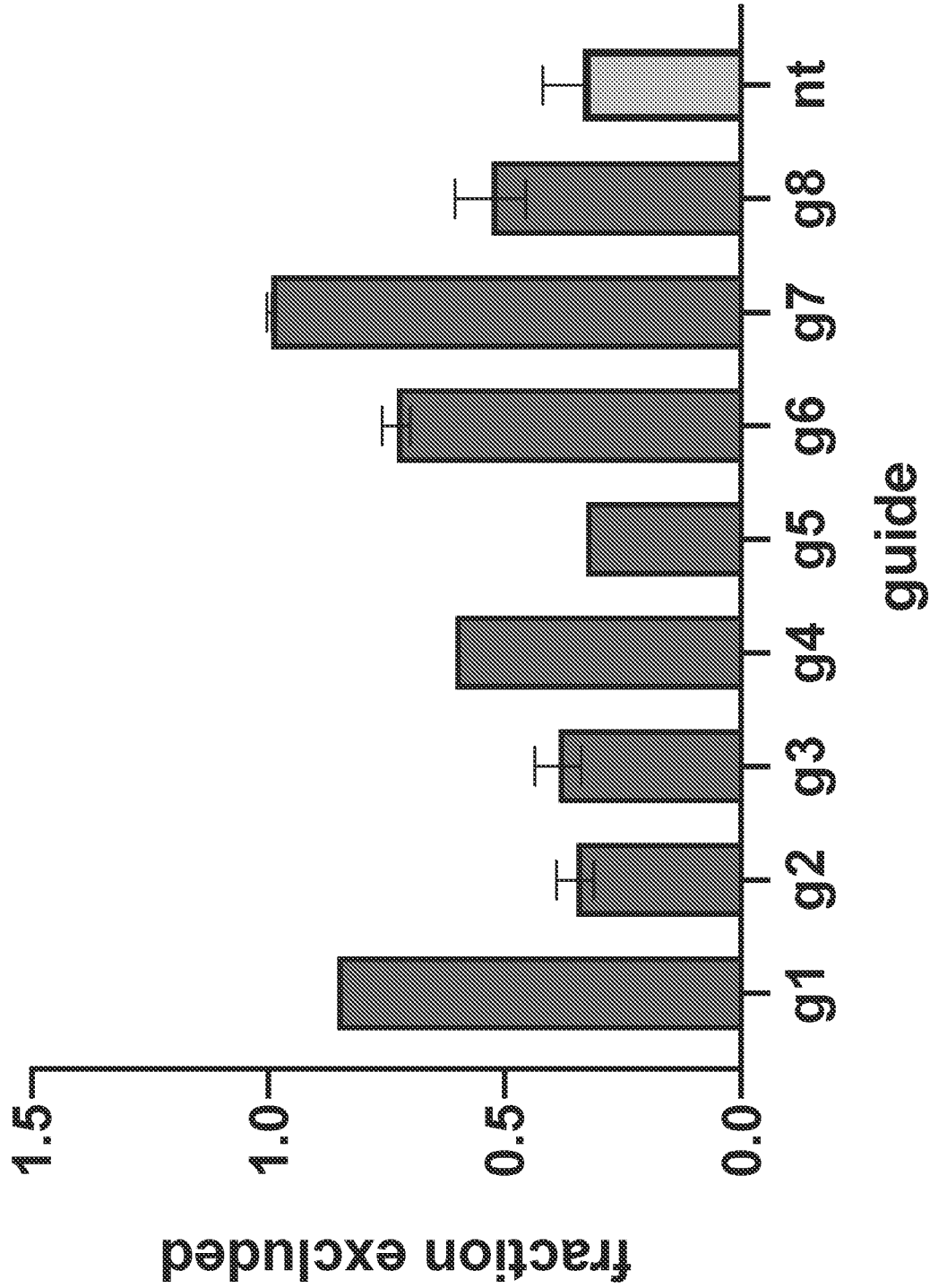


FIG. 3E

FIG. 3D

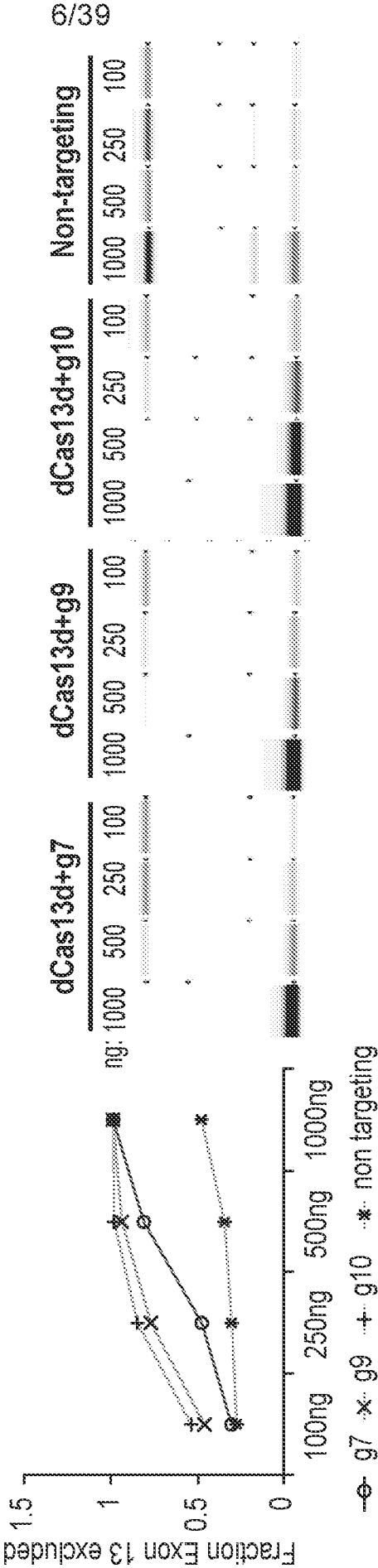


FIG. 4A

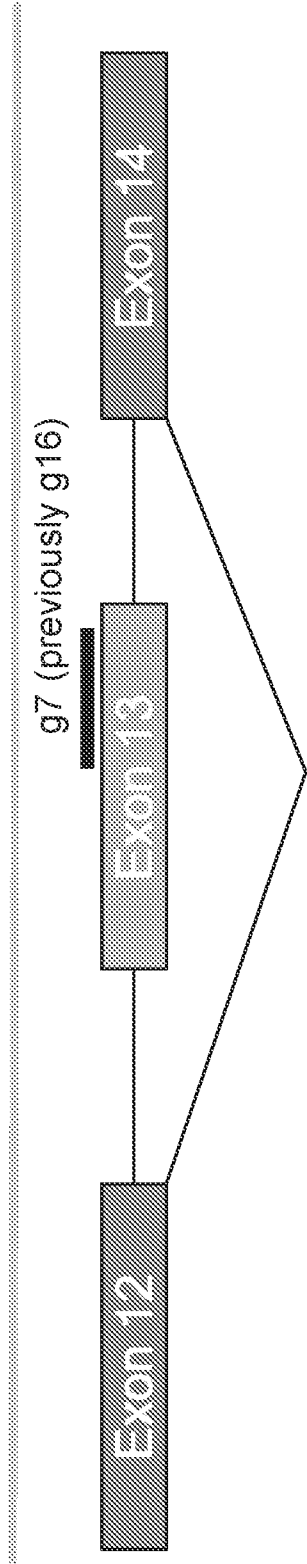


FIG. 4B

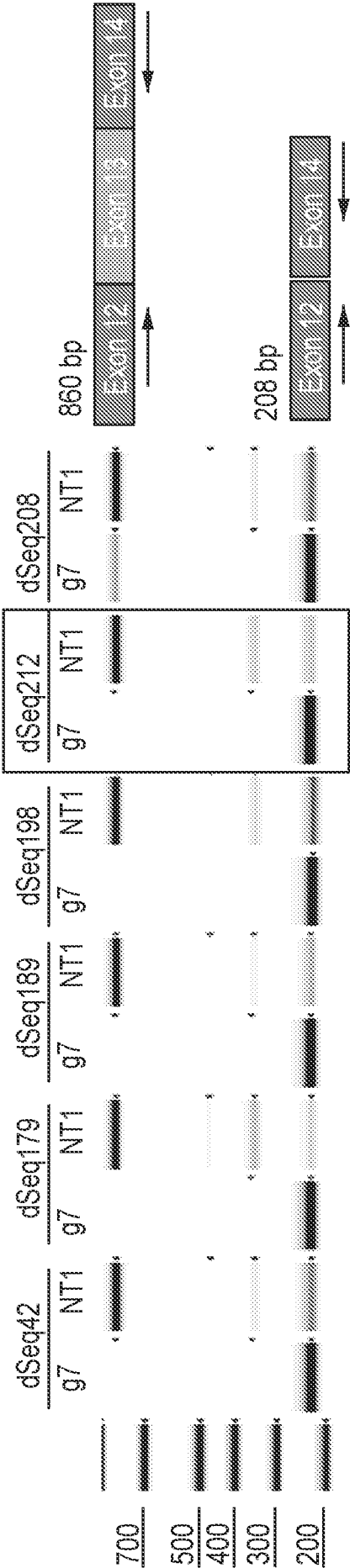


FIG. 5A

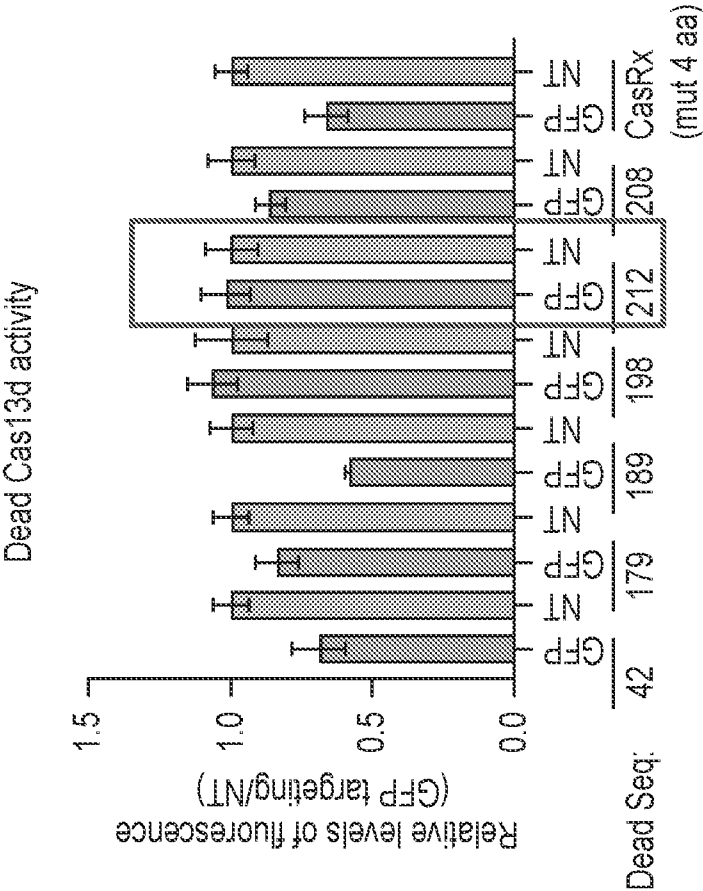
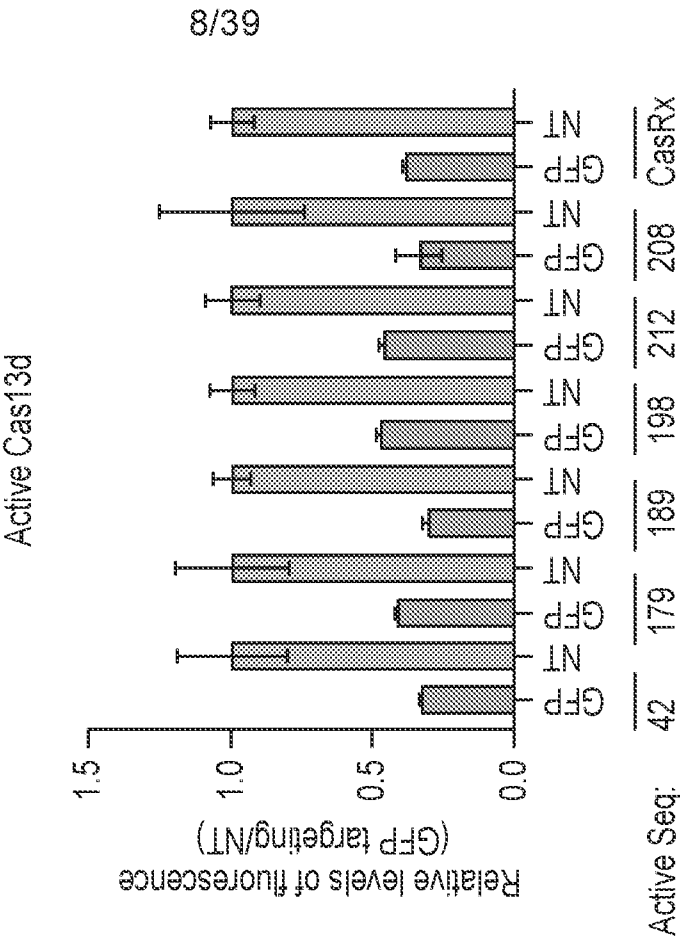


FIG. 5B



8/39

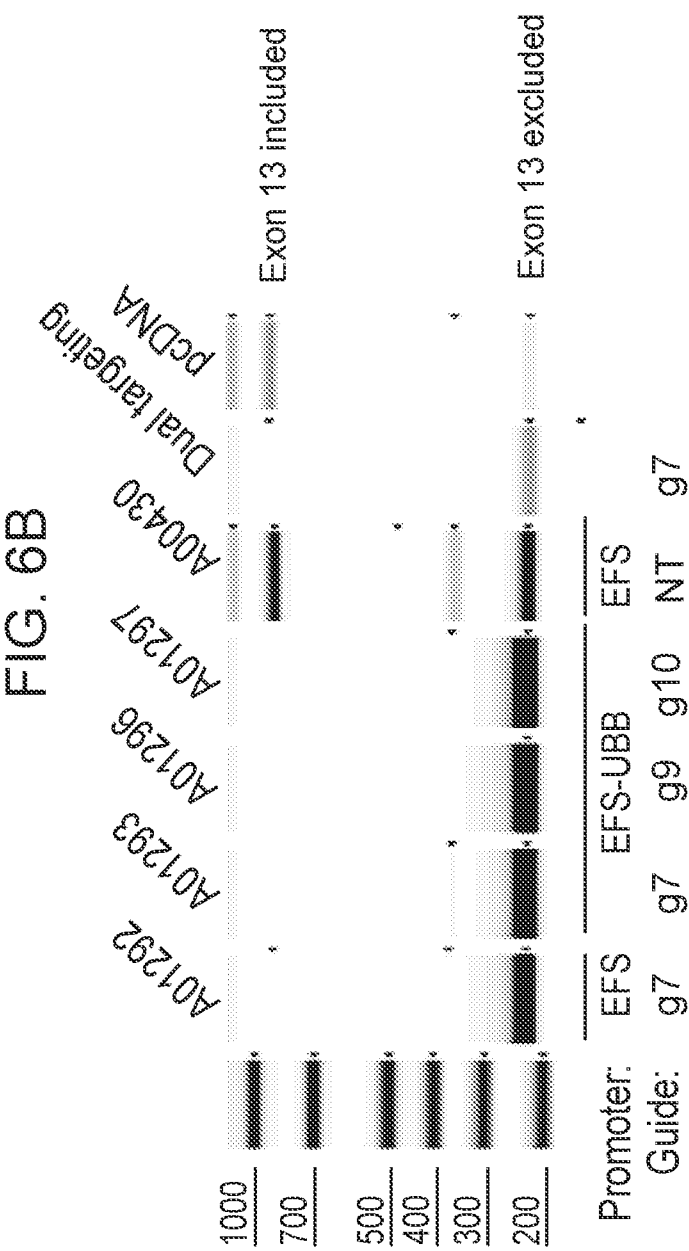
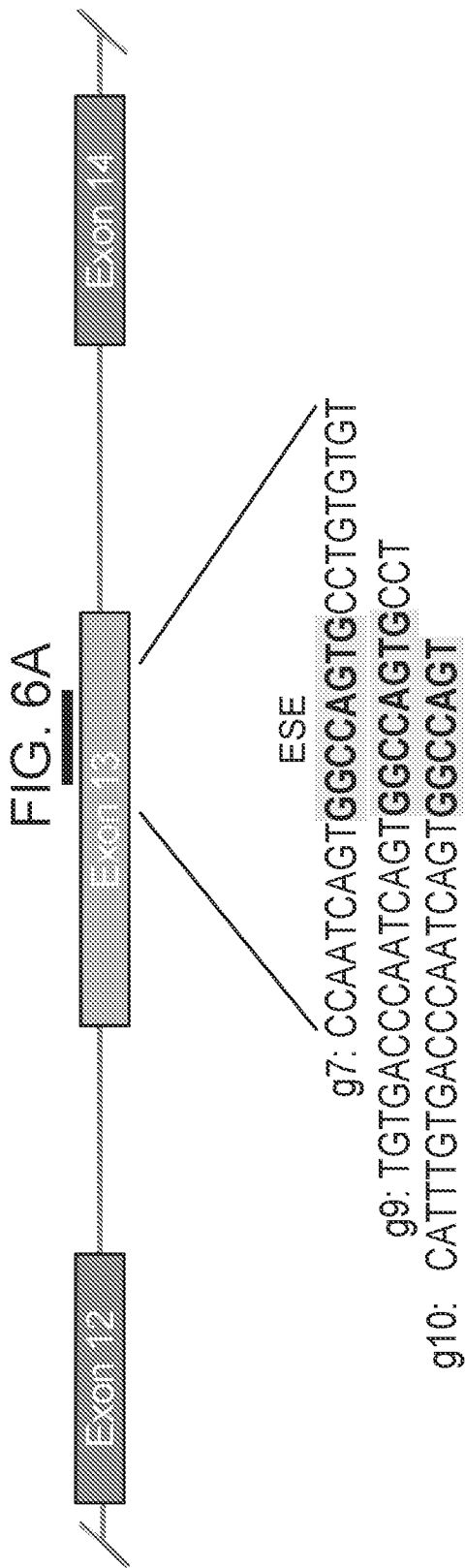


FIG. 7A

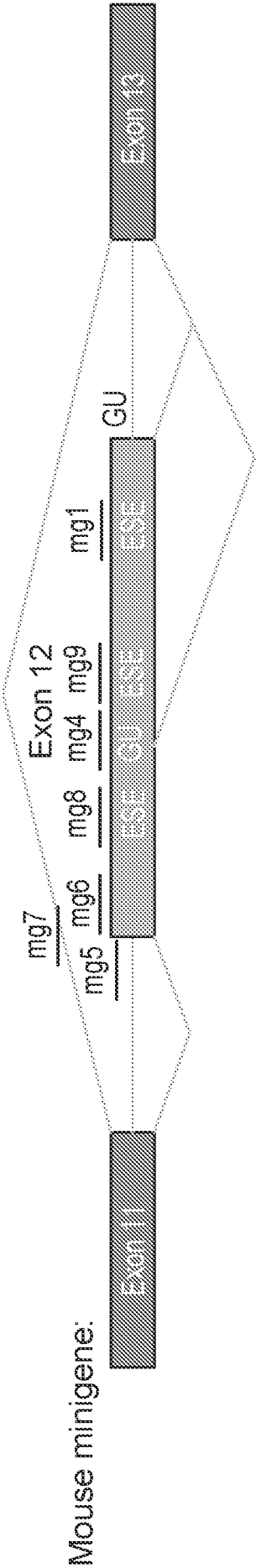
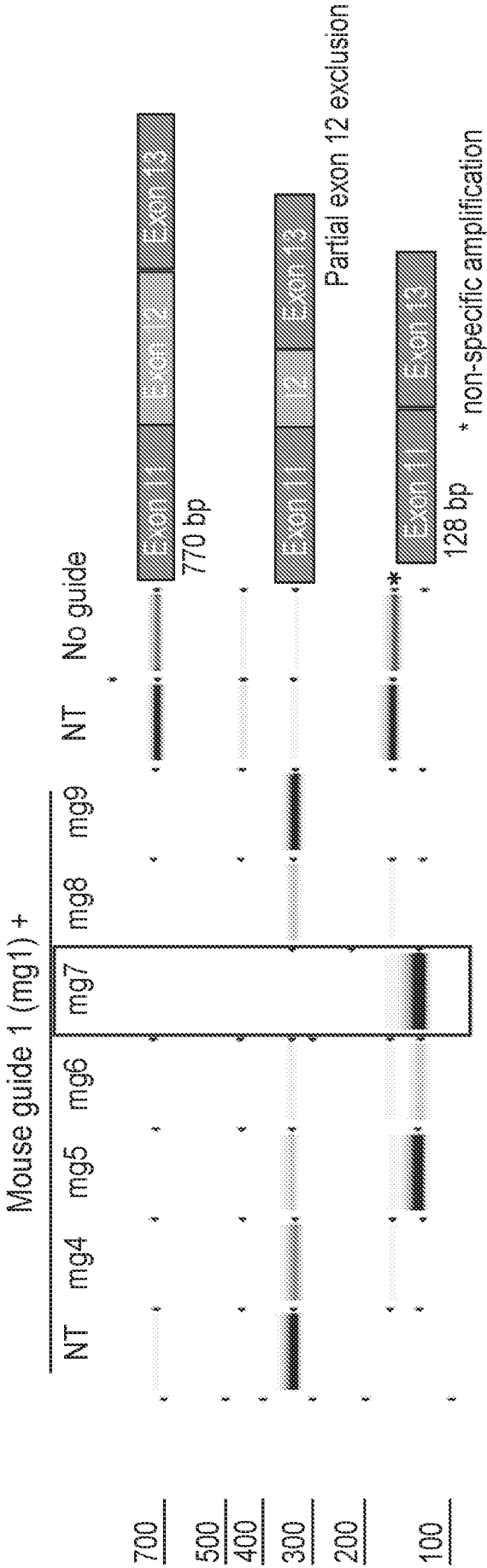
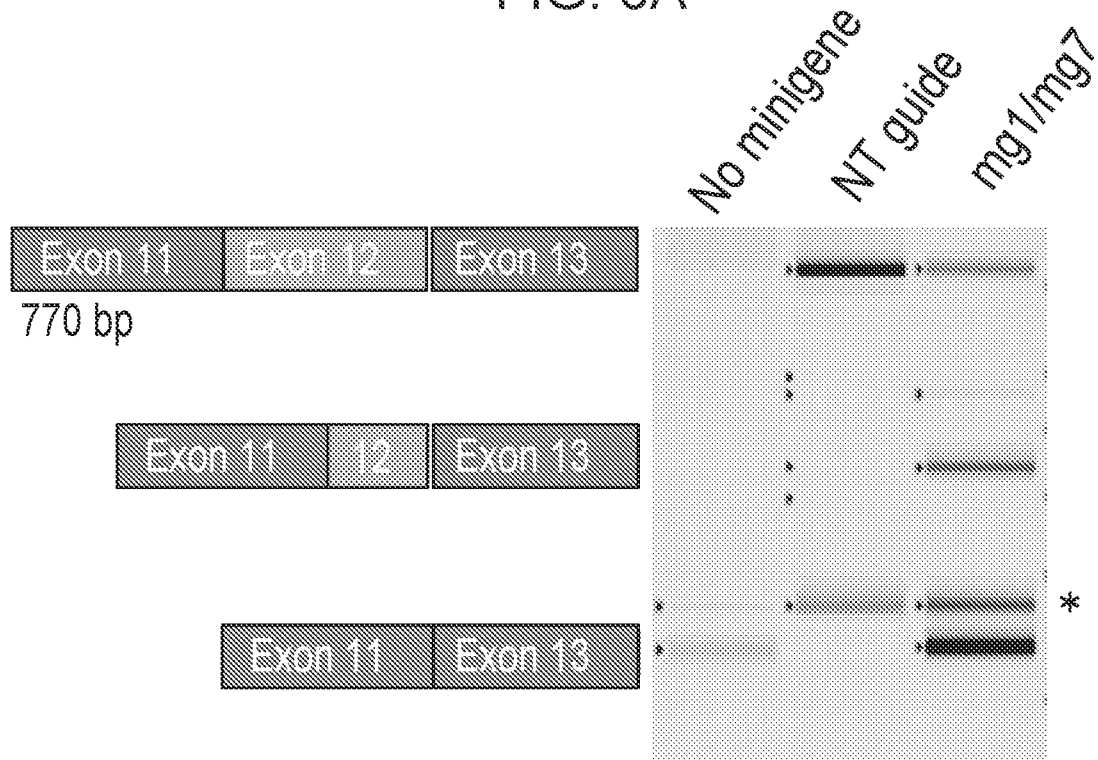


FIG. 7B



11/39

FIG. 8A



\*non-specific amplification from minigene

FIG. 8B

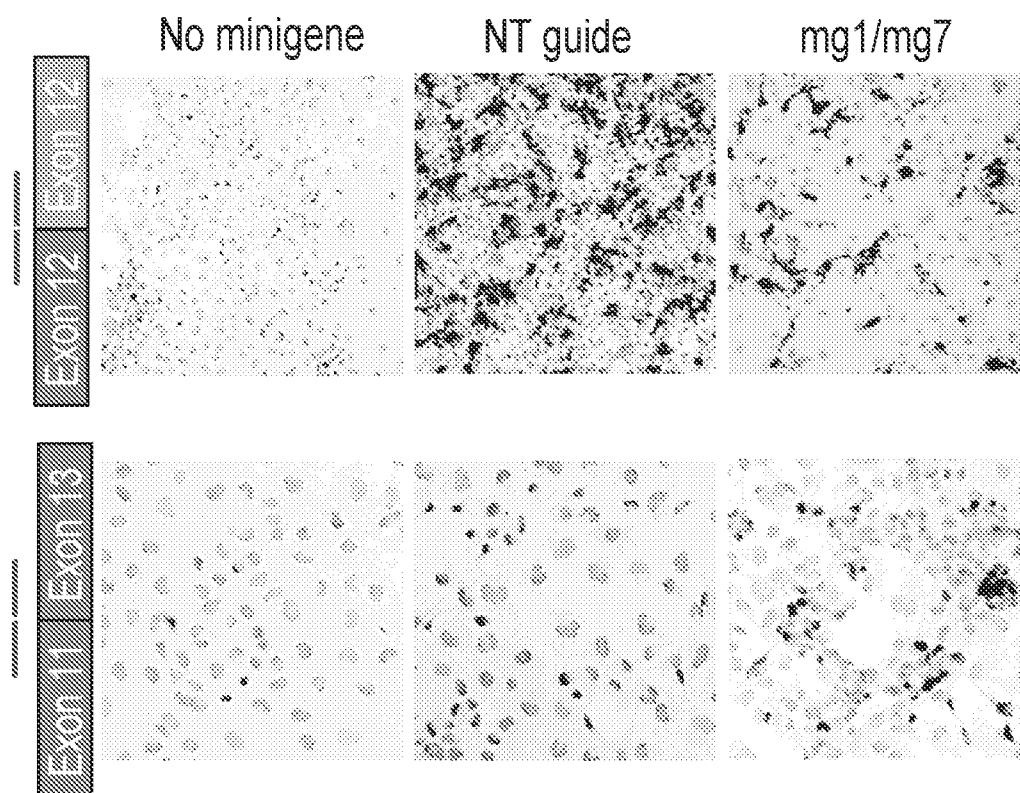


FIG. 9A

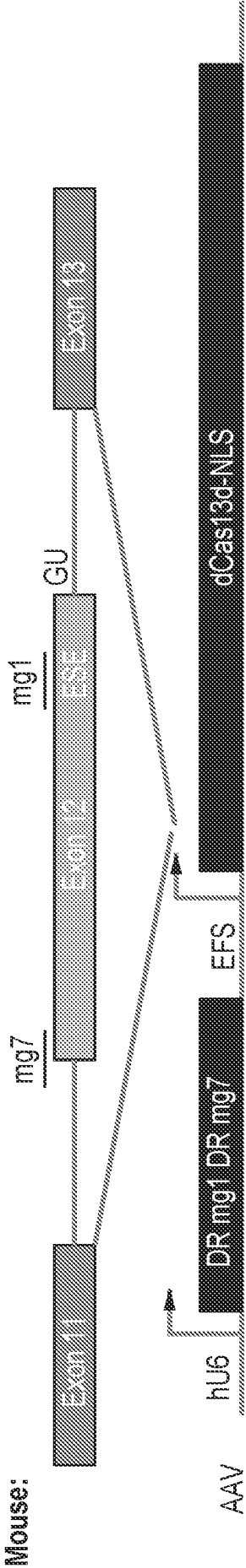


FIG. 9B

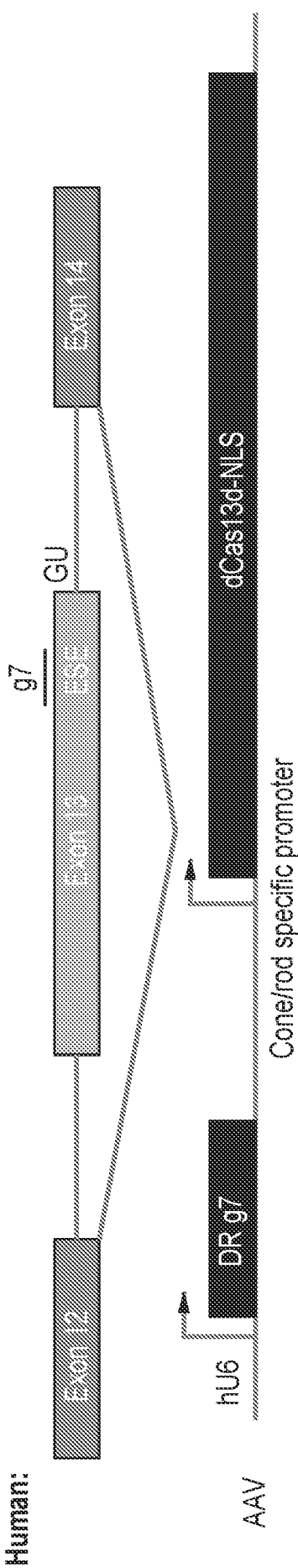


FIG. 10

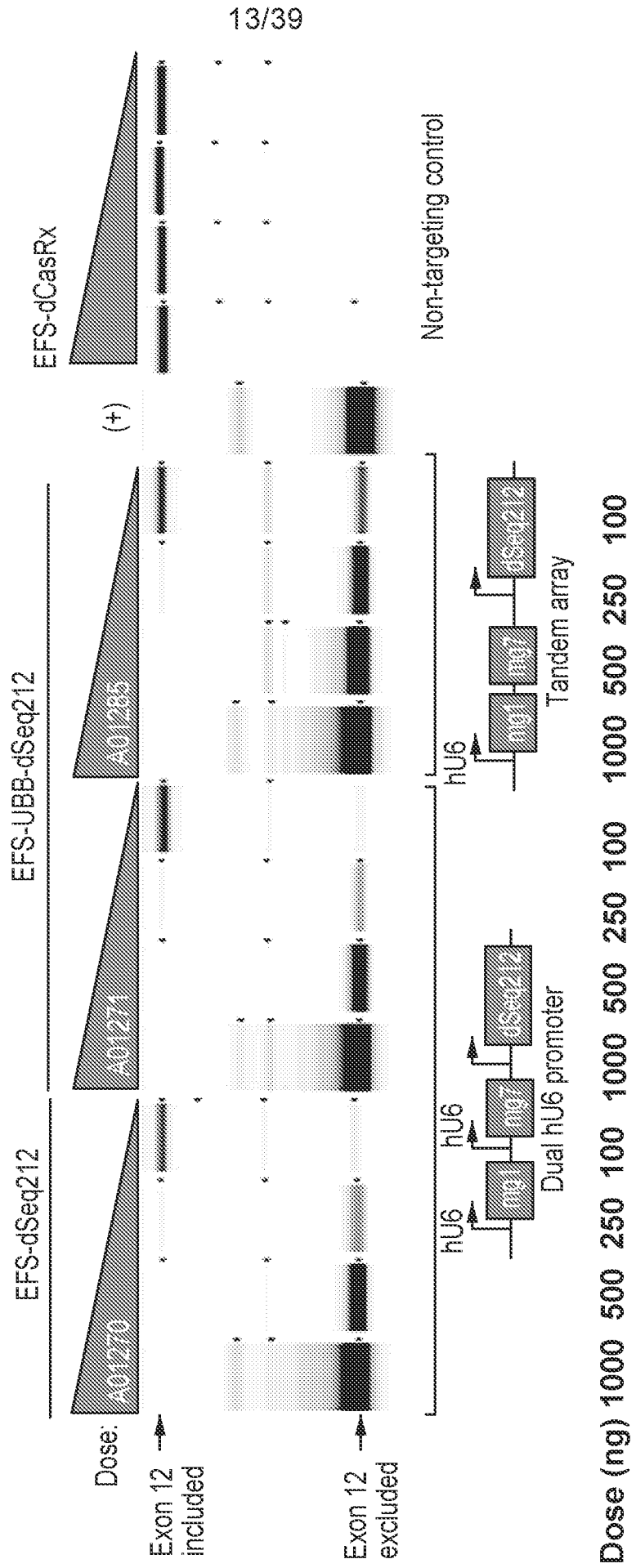


FIG. 11A

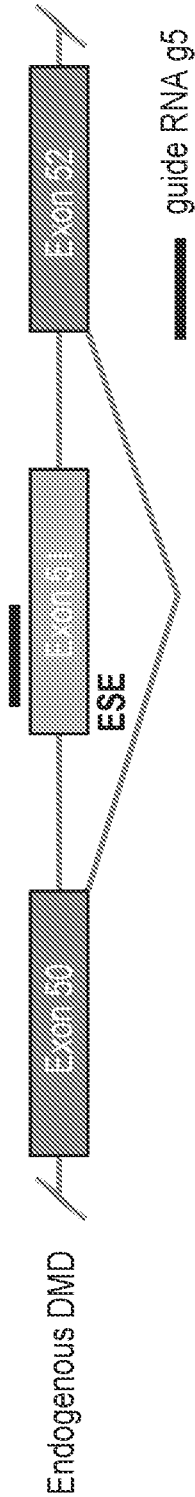


FIG. 11B

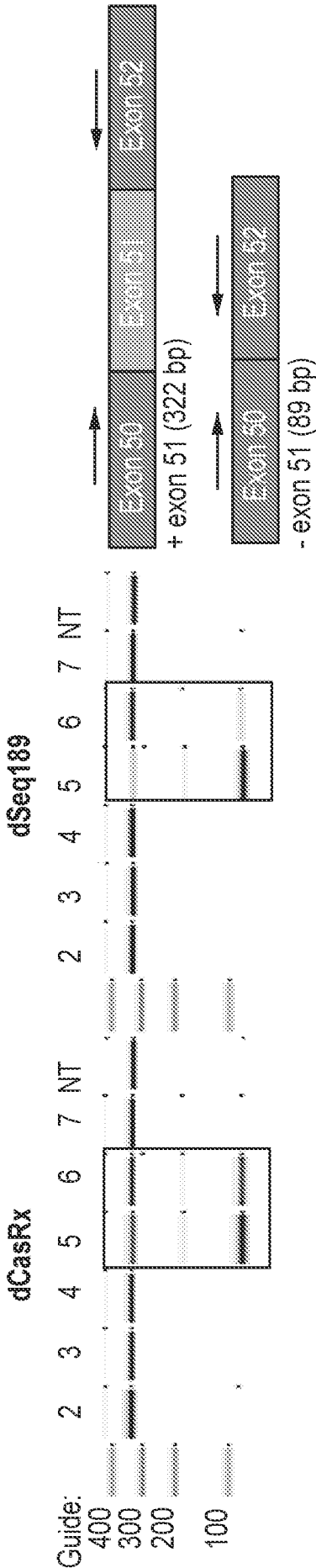
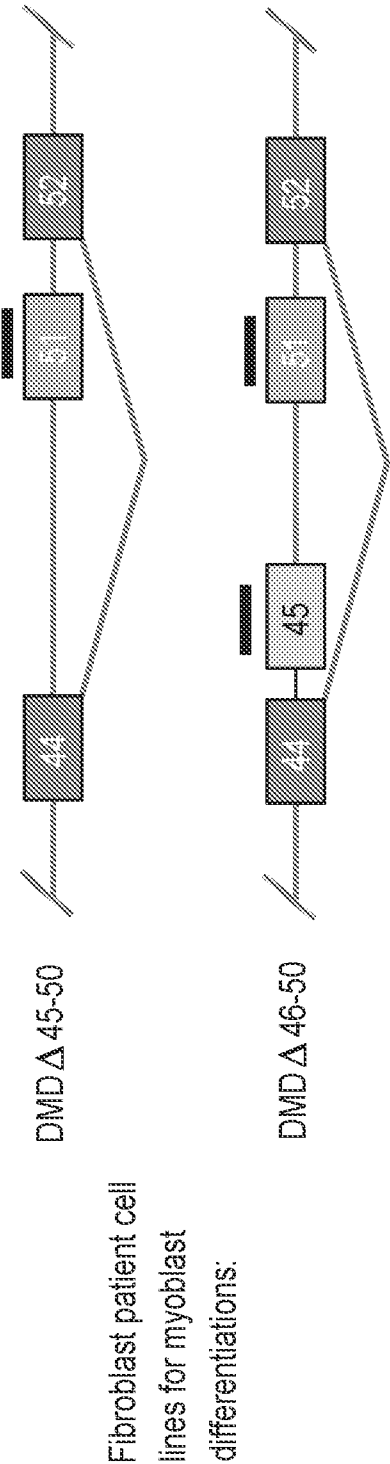
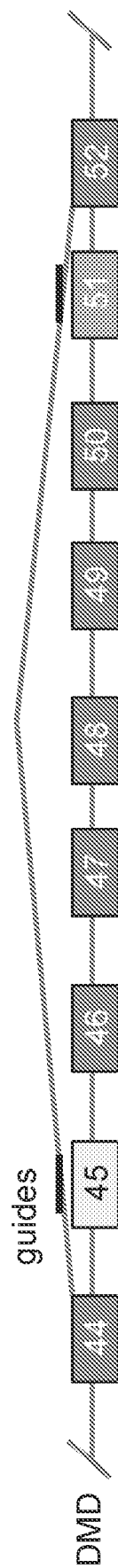


FIG. 12



13A  
G<sup>5</sup>  
L



ம  
ம  
த  
த  
த  
த

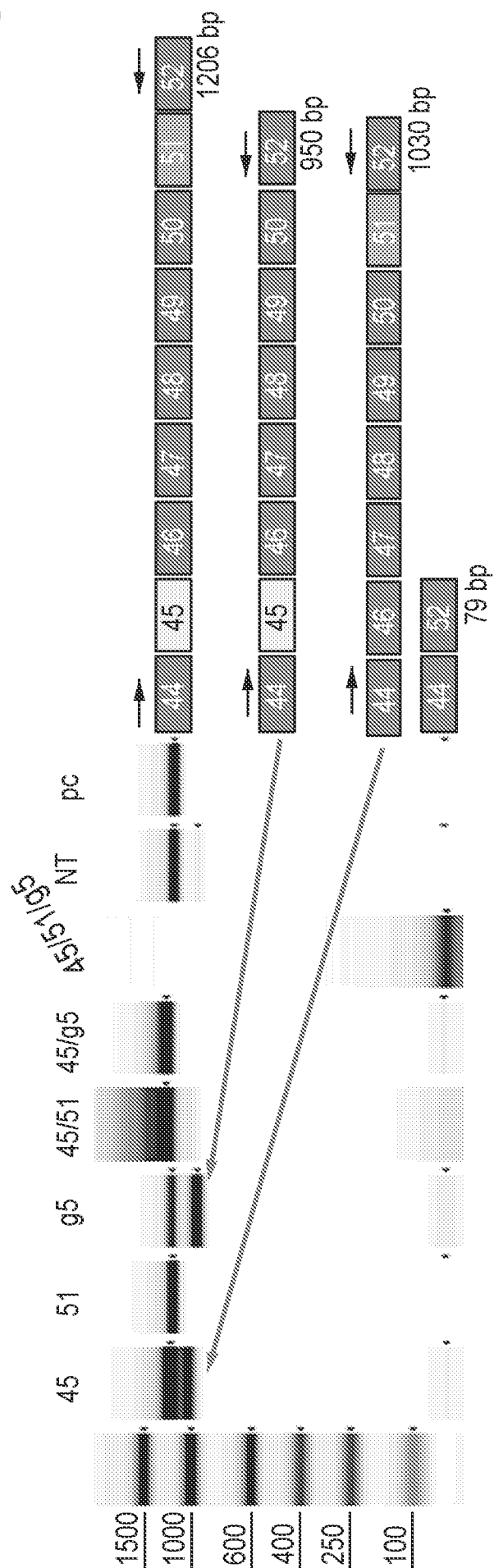


FIG. 14

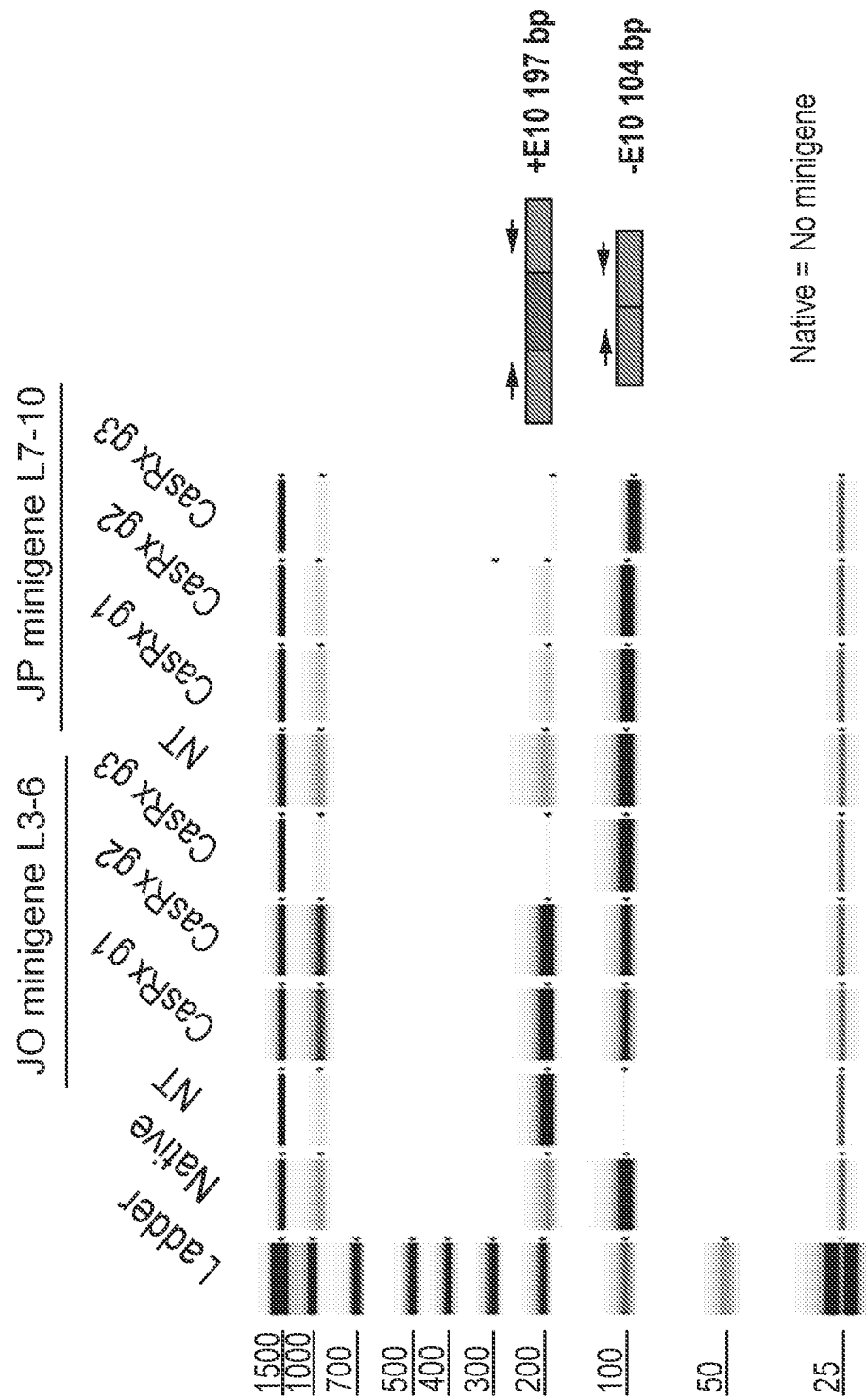


FIG. 15

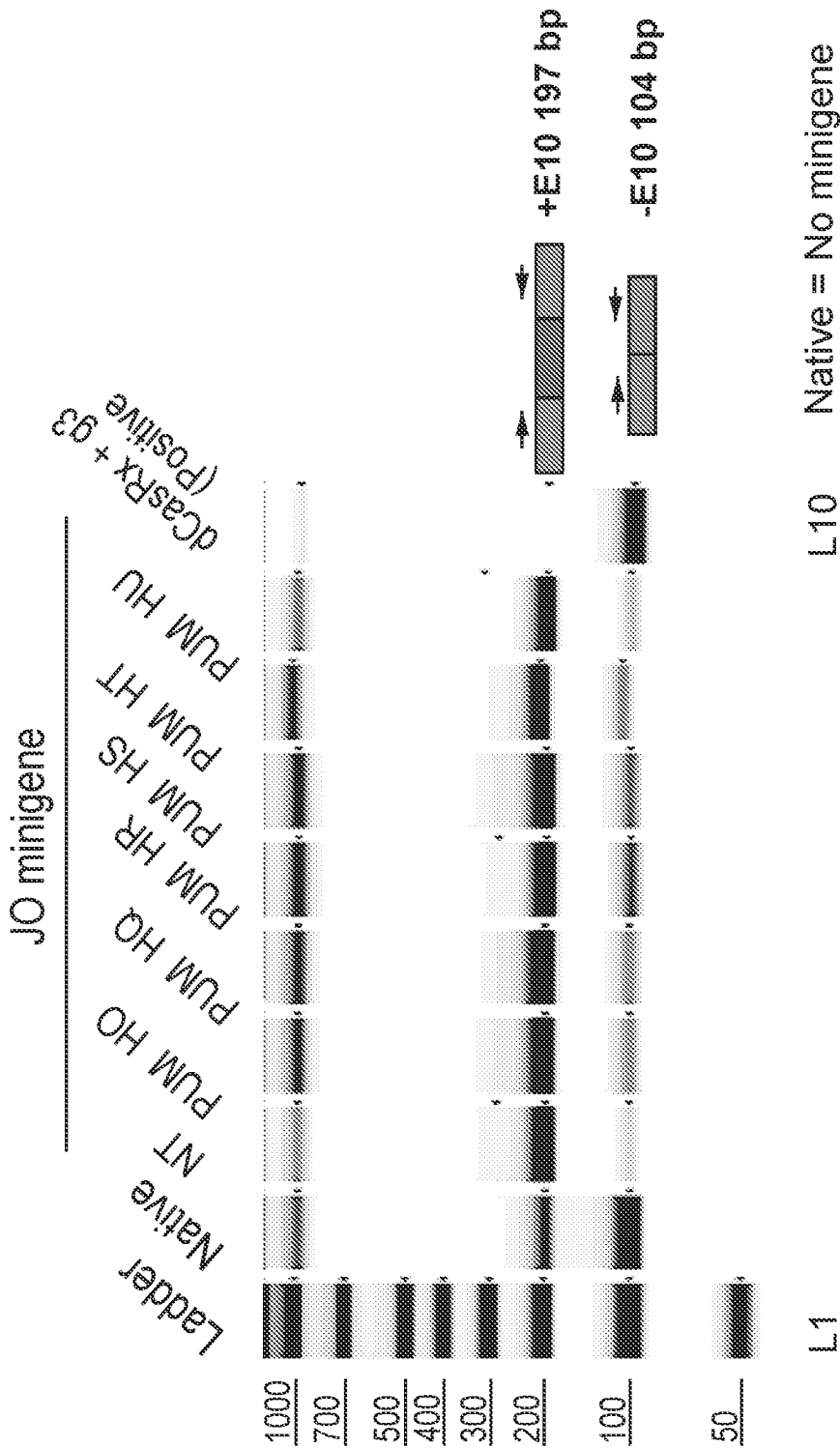
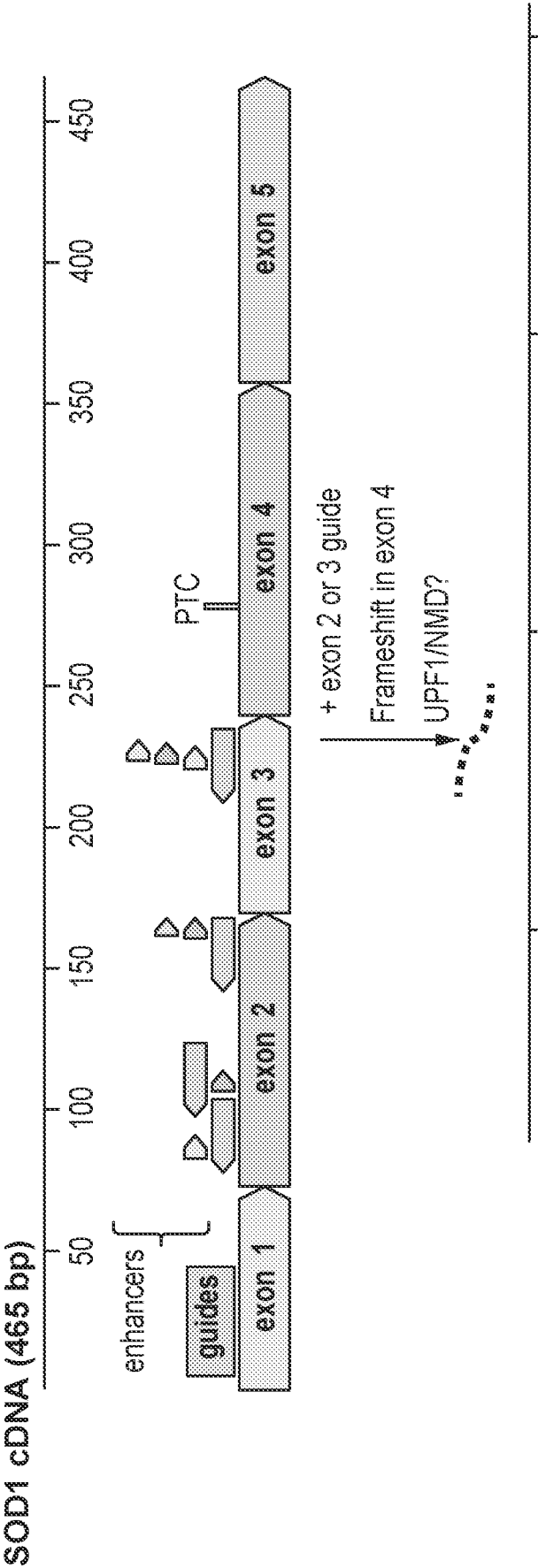
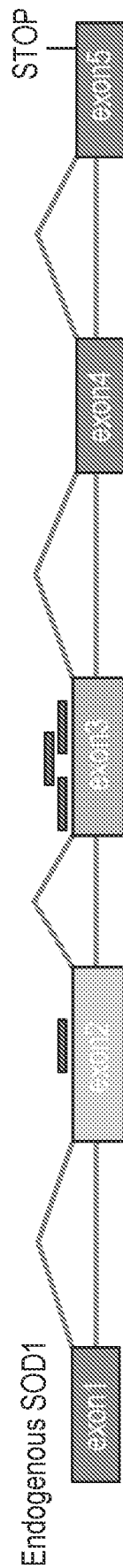


FIG. 16A





100

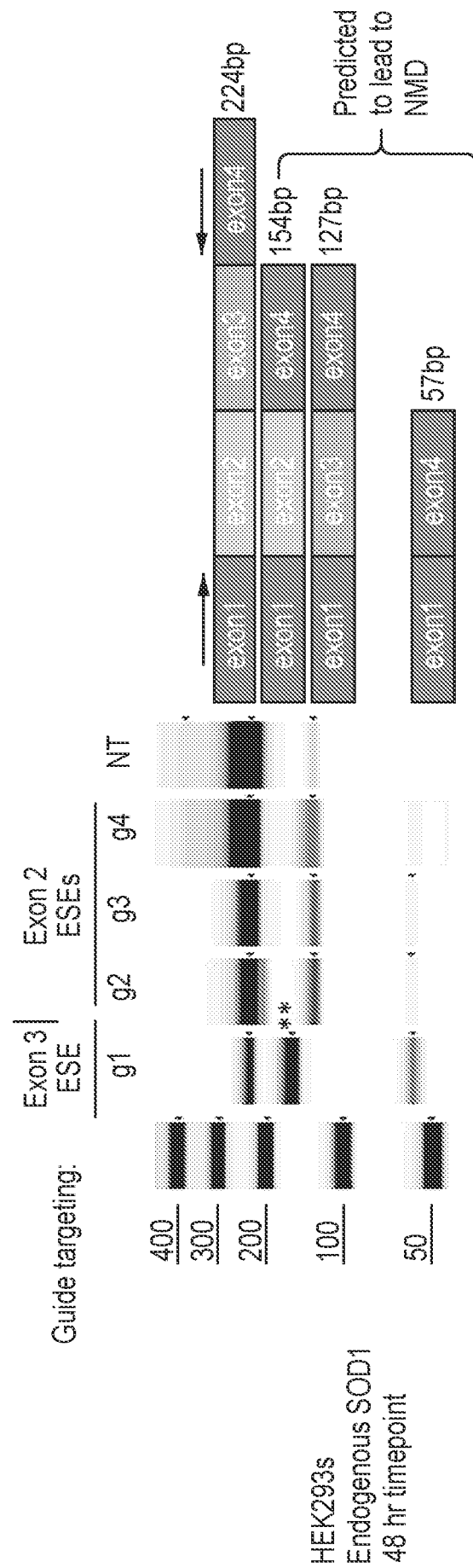


FIG. 17

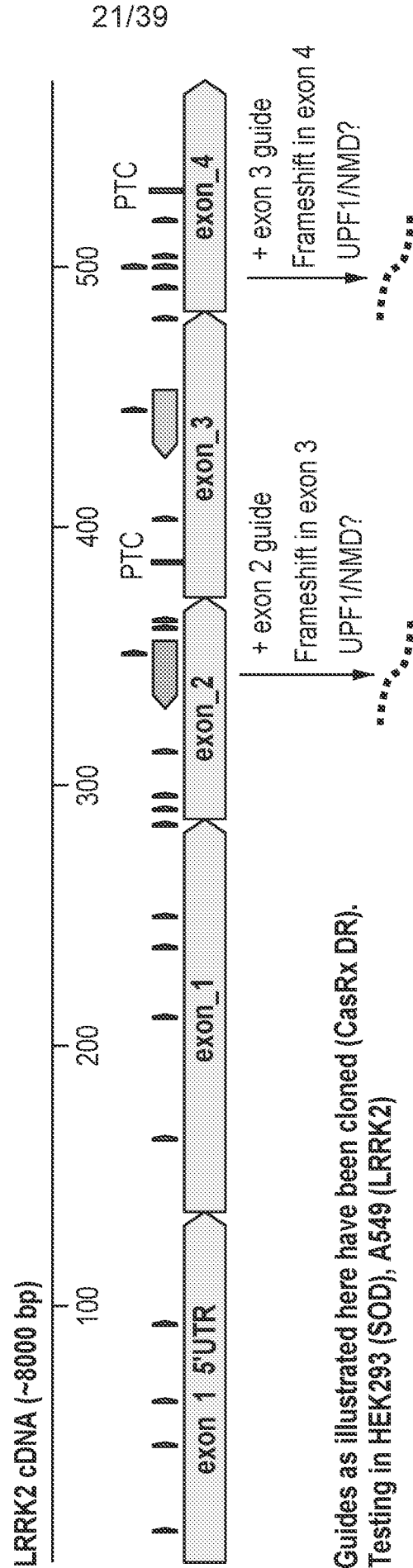


FIG. 18A

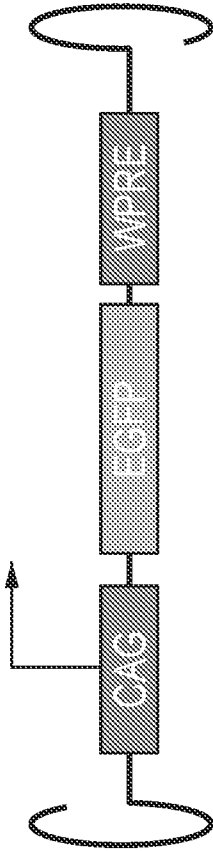
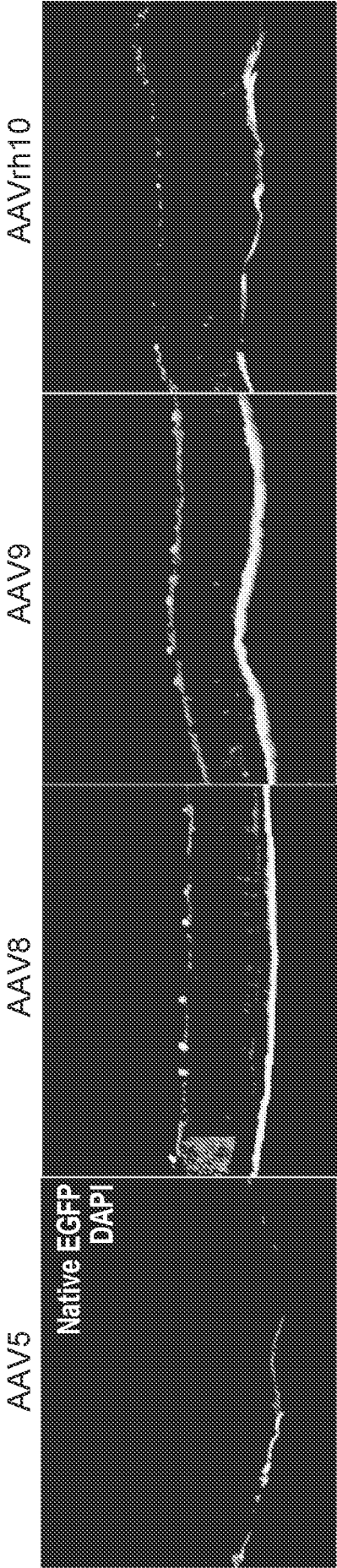


FIG. 18B



22/39

FIG. 18C

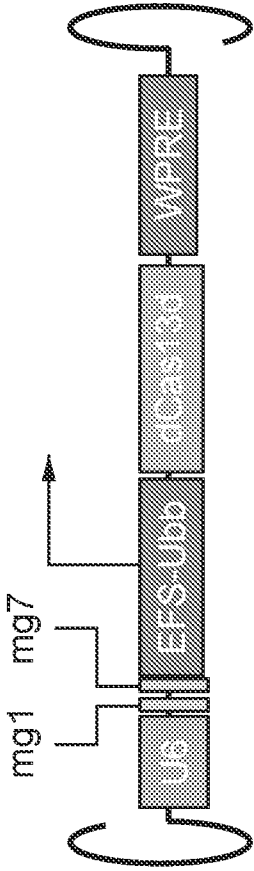
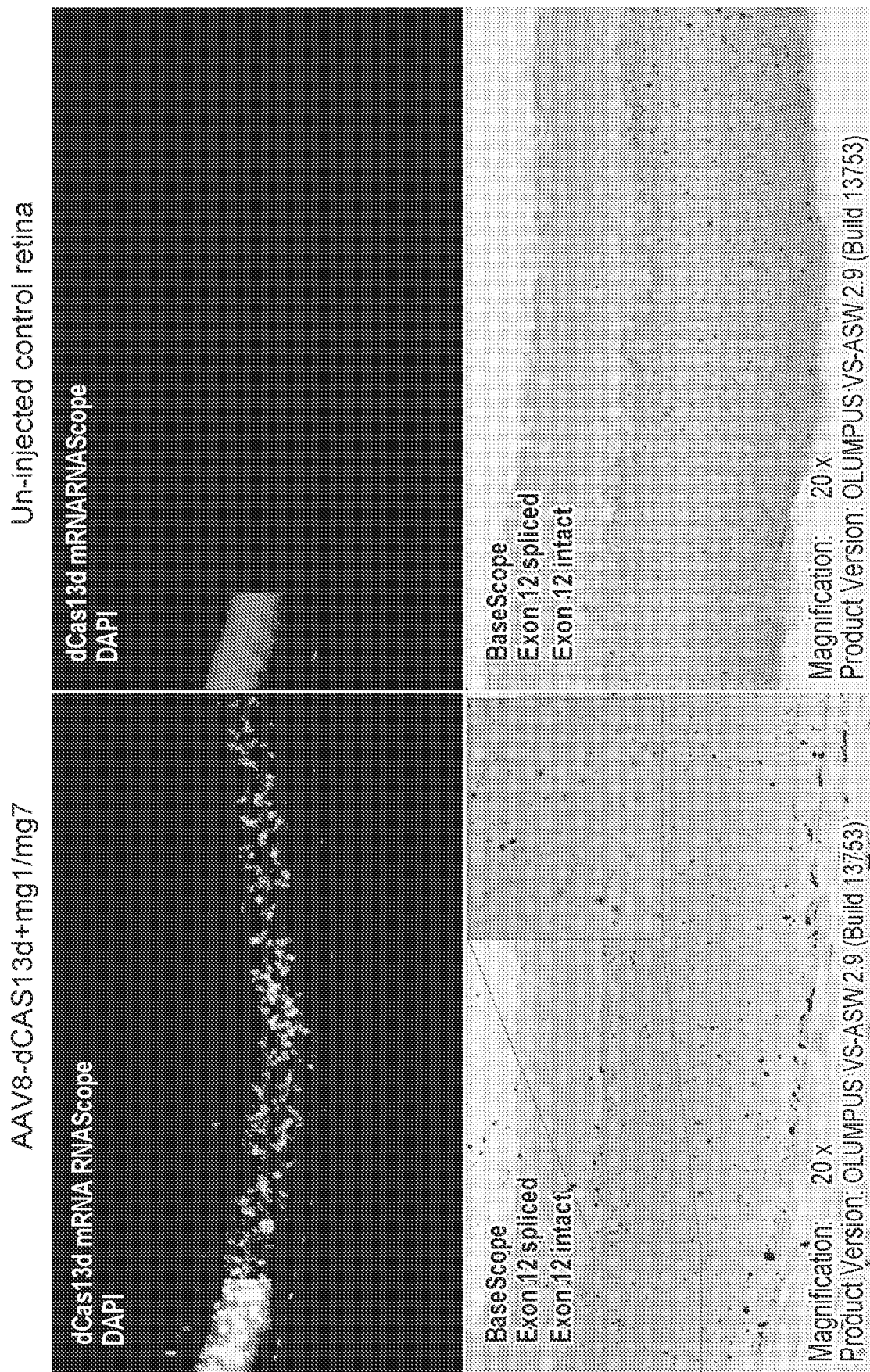


FIG. 18D



24/39

FIG. 18E

Exon 12 skipping in treated vs. control  
photoreceptors

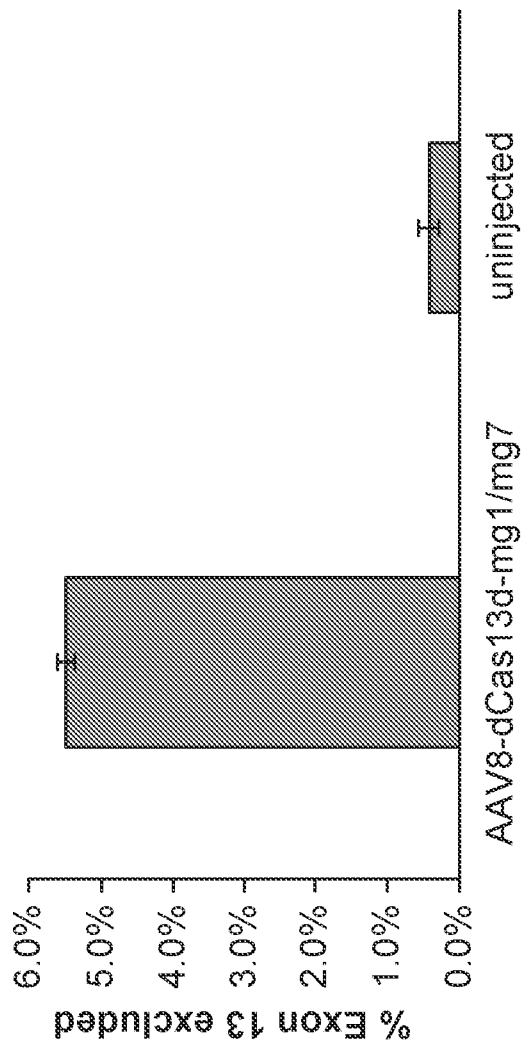


FIG. 19A

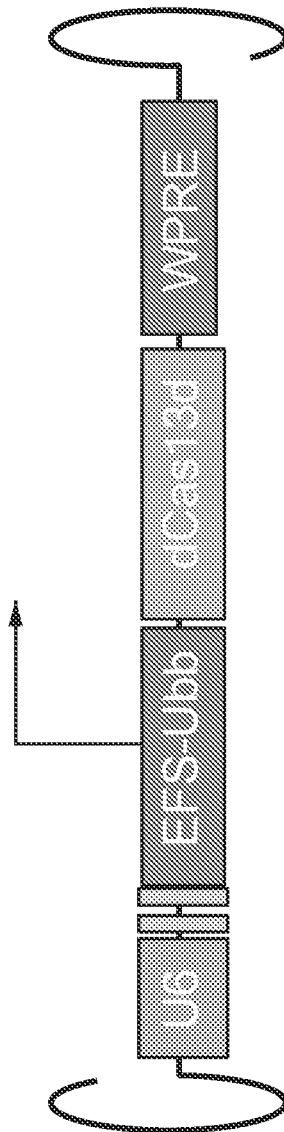
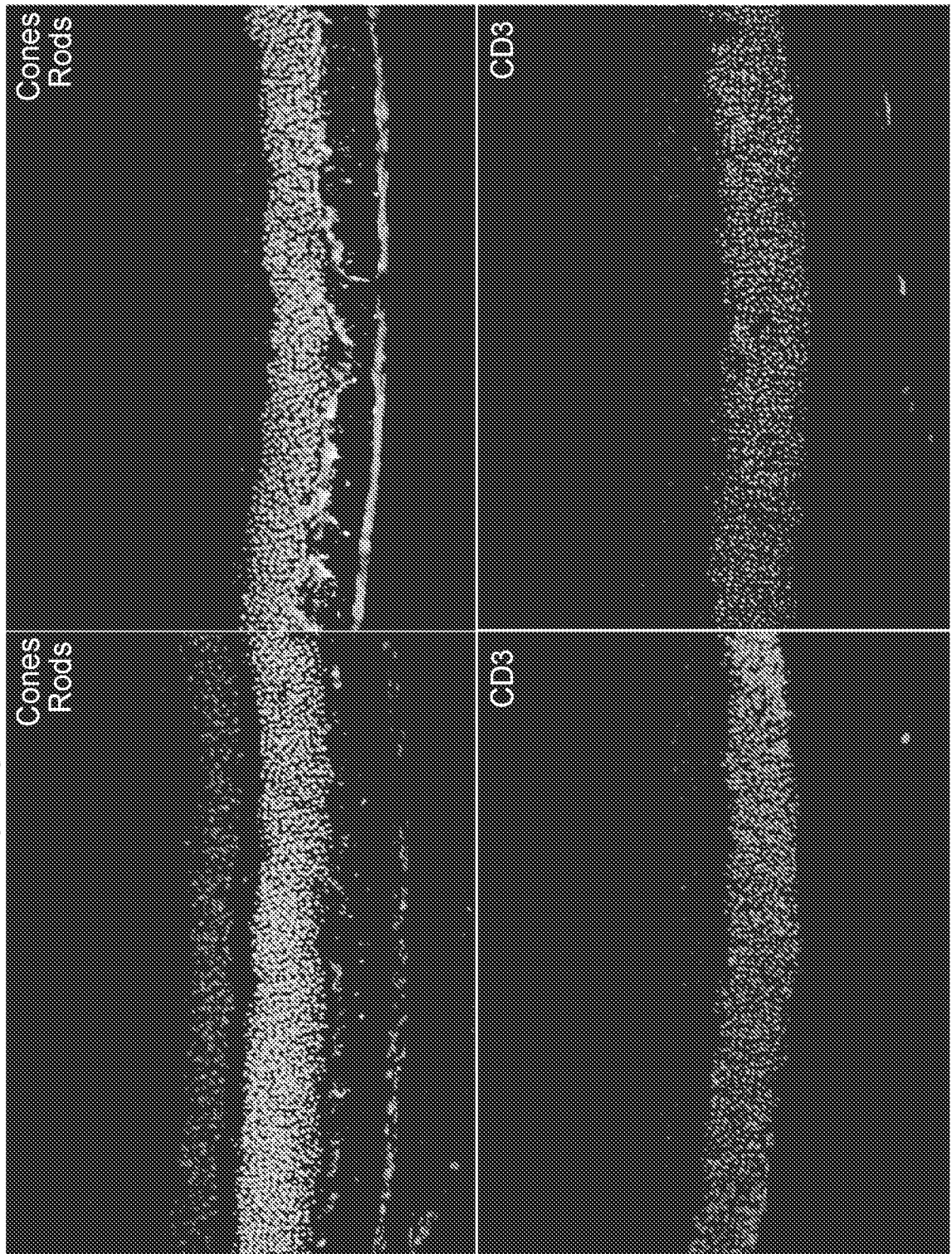


FIG. 19B

AAV8-dCASP13d+mg1/mg7

Un-injected control retina



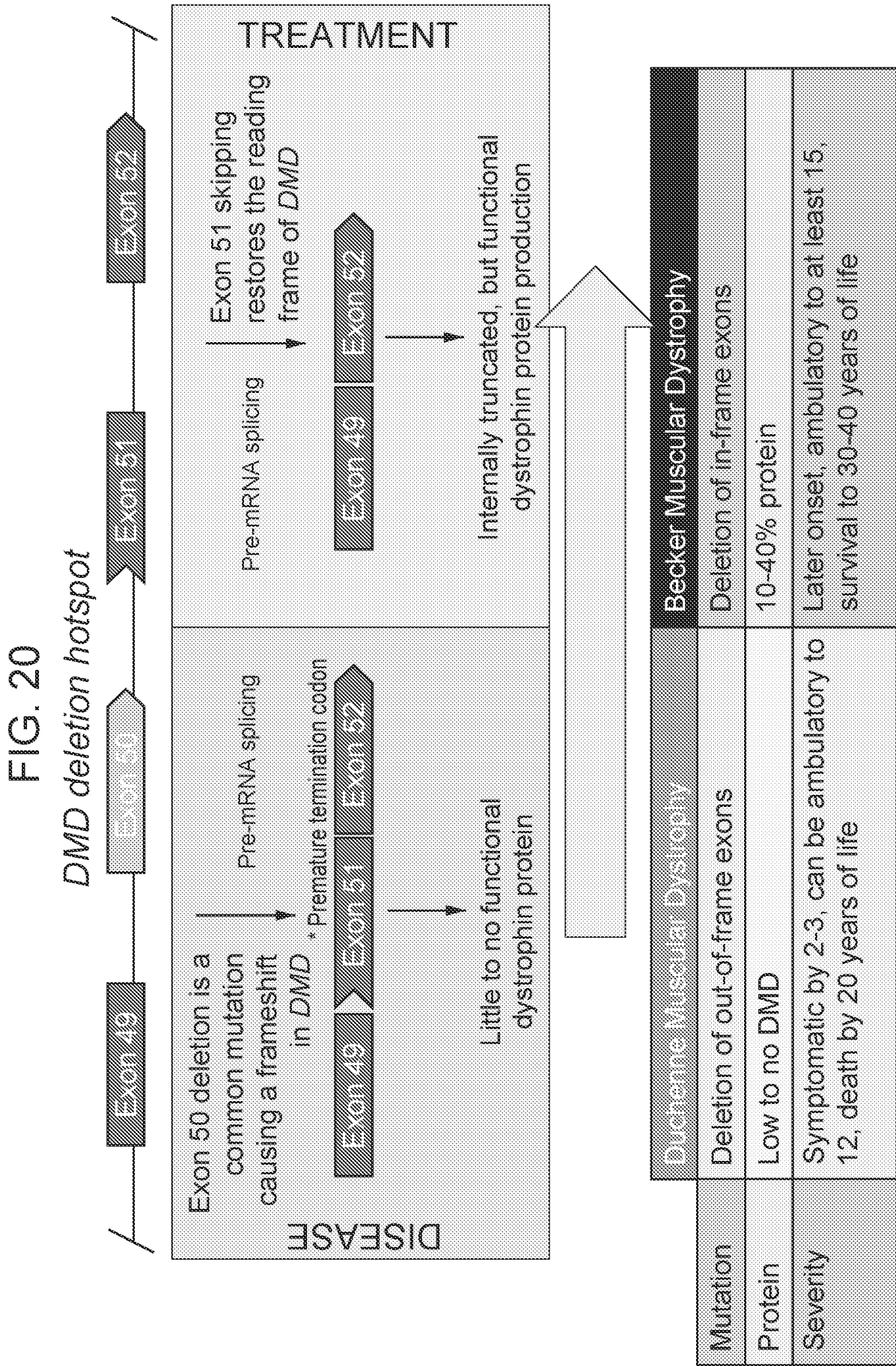


FIG. 21

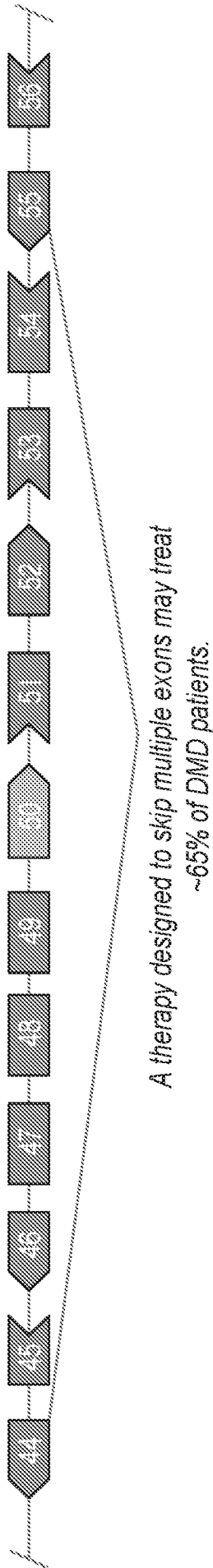
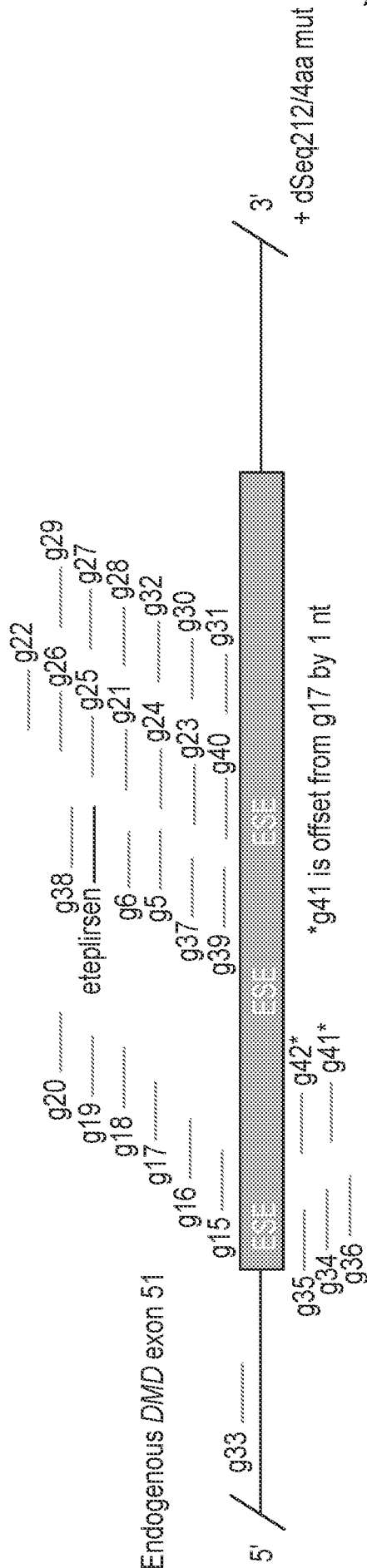


FIG. 22A



29/39

FIG. 22B

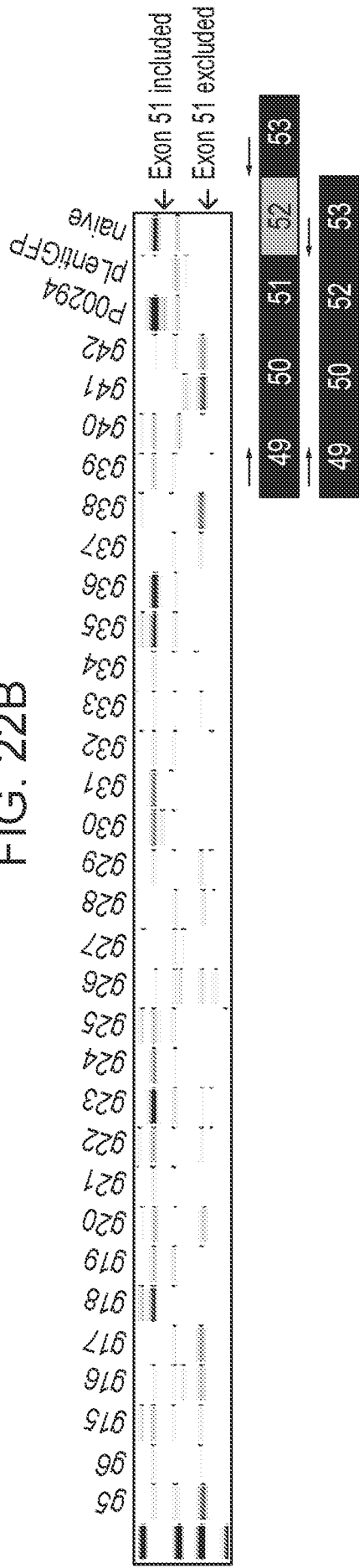


FIG. 22C

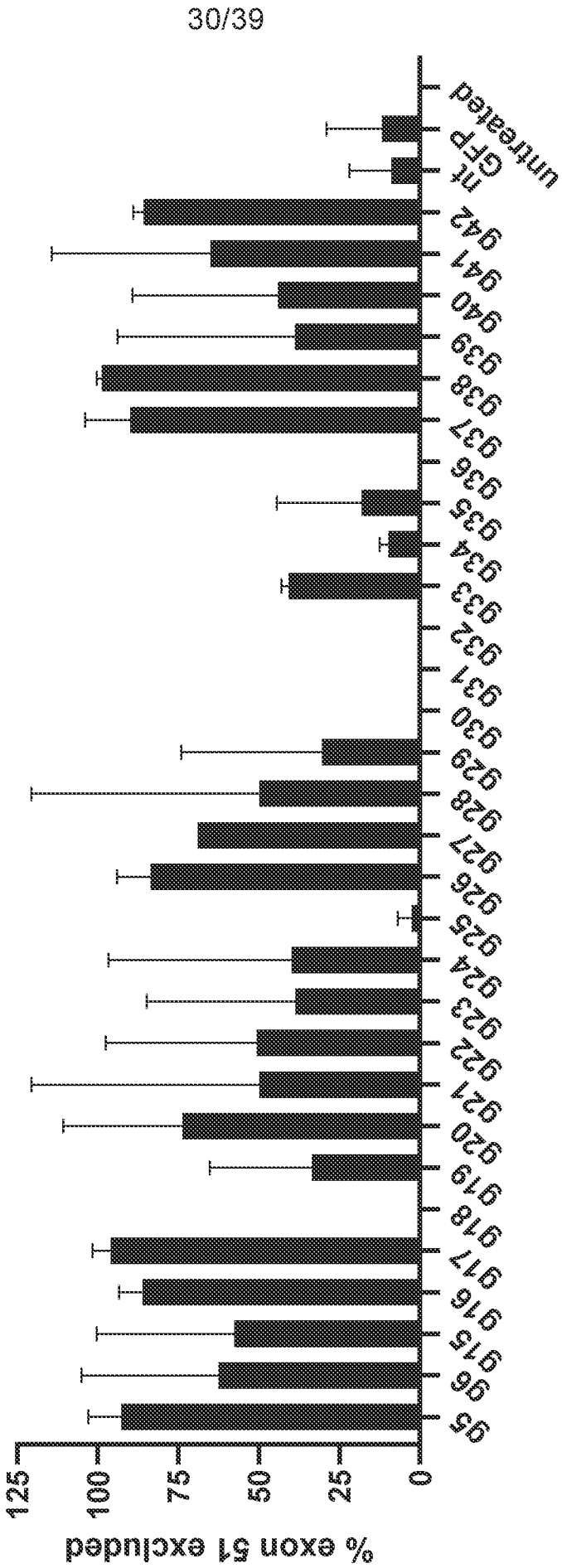


FIG. 23A

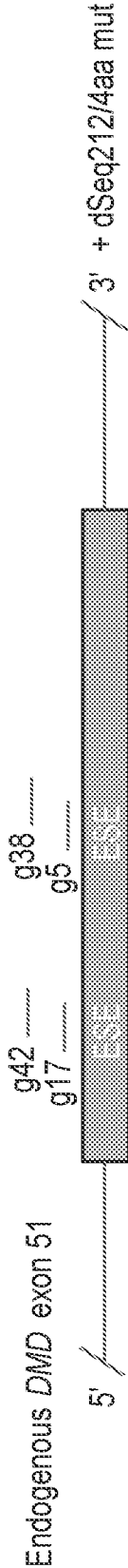
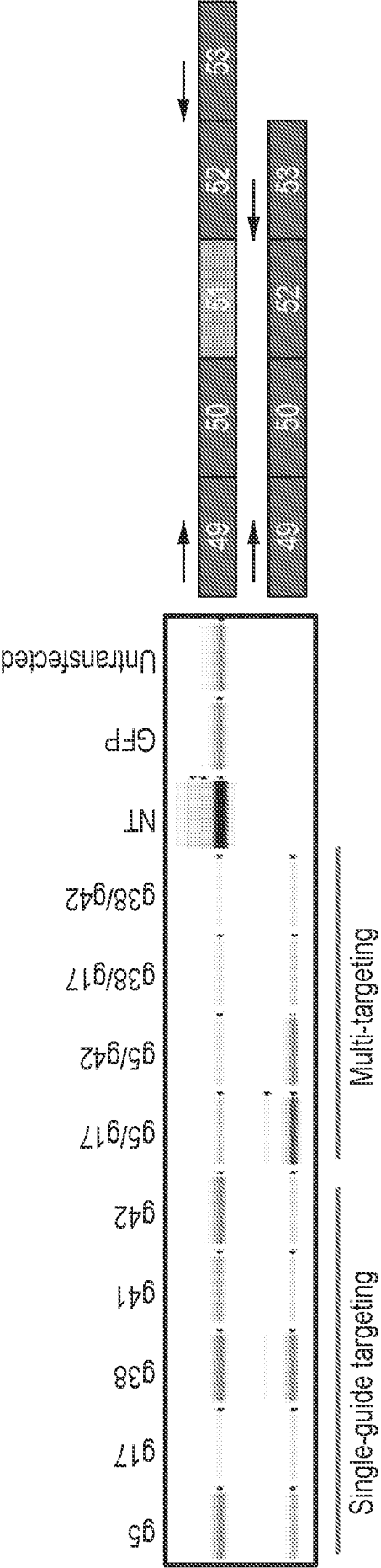


FIG. 23B



32/39

FIG. 23C

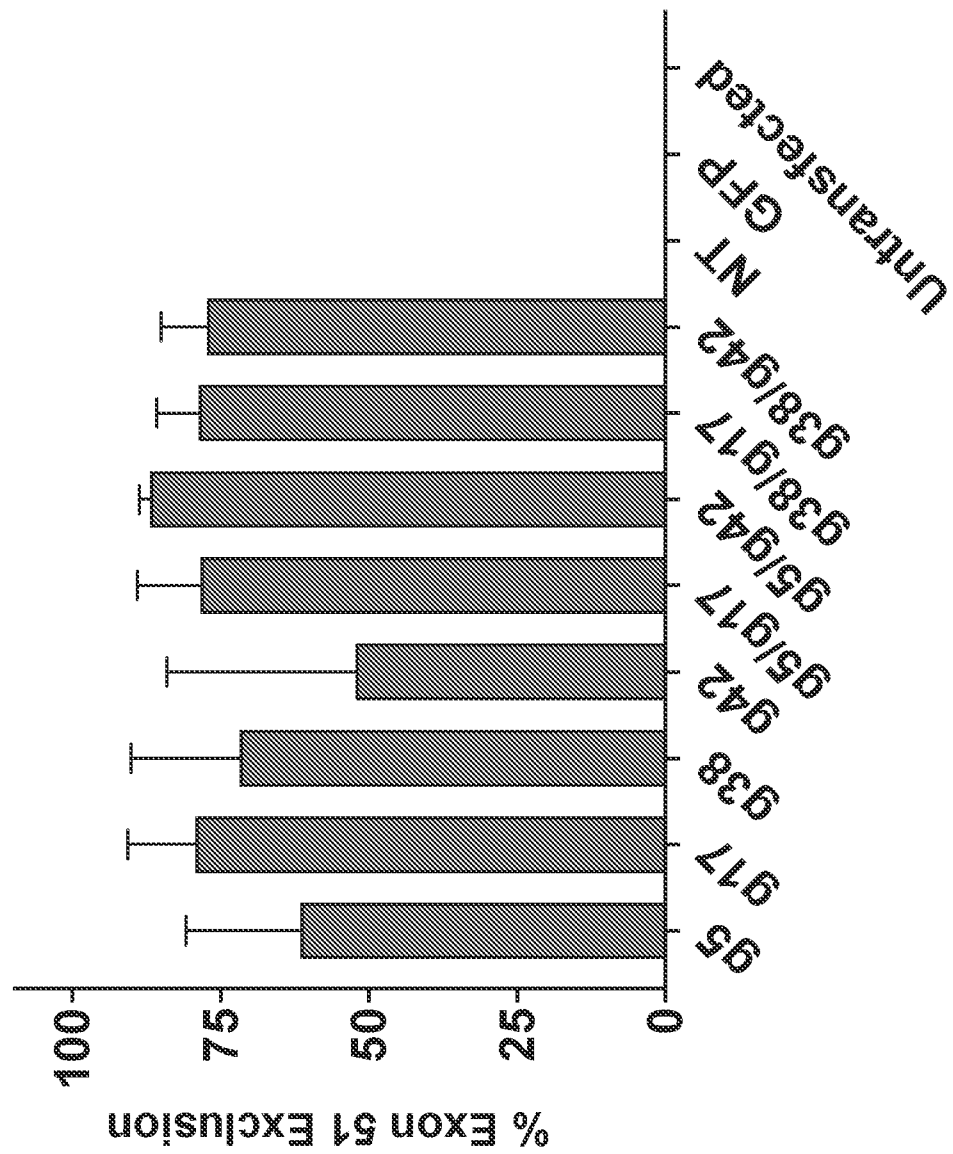
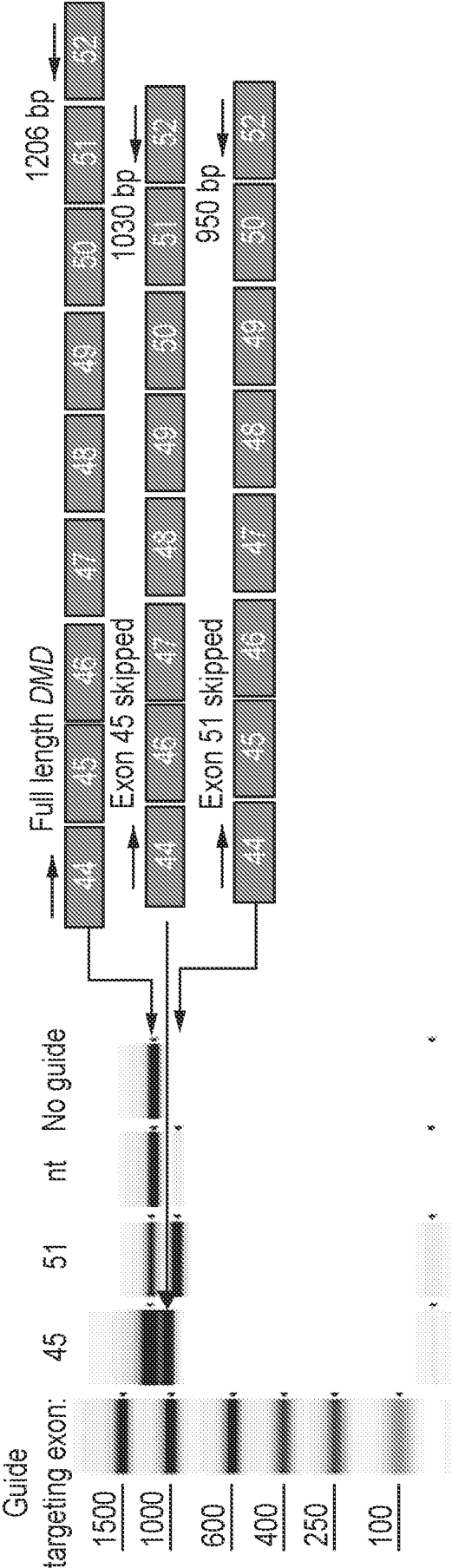
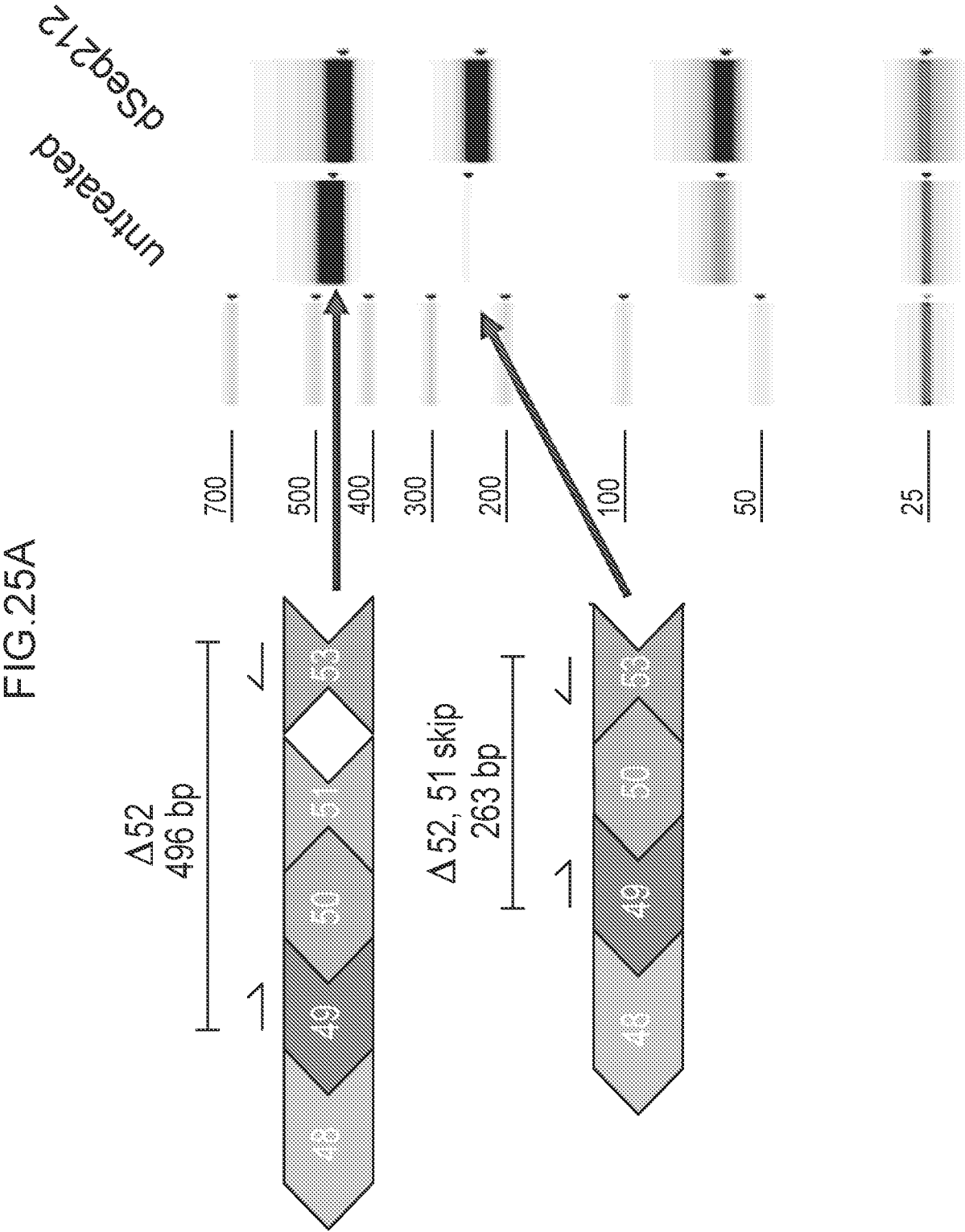


FIG. 24





35/39

FIG. 25B

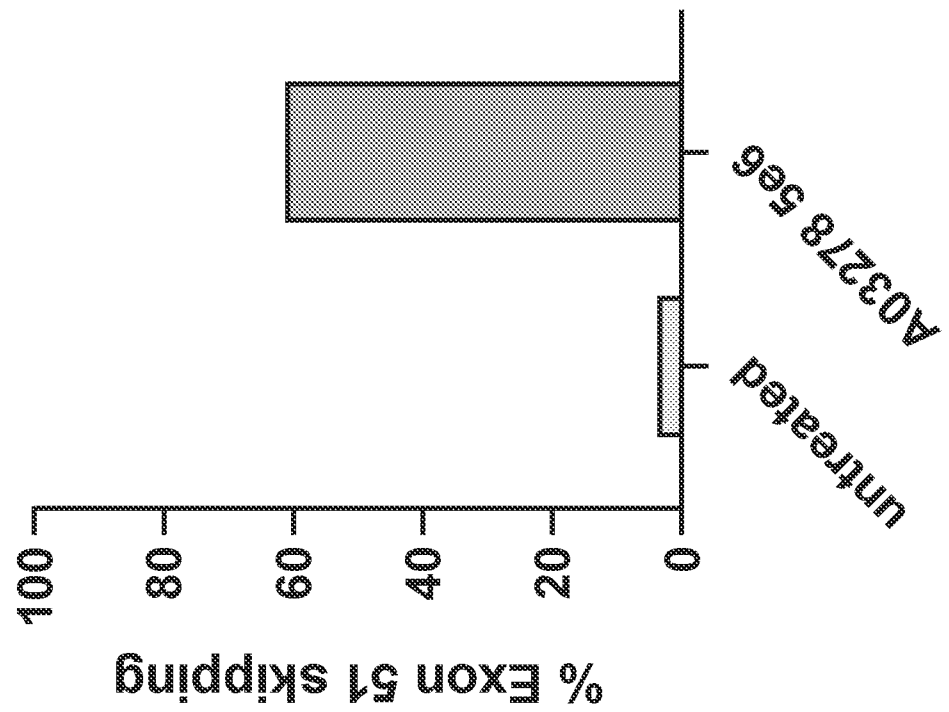


FIG. 25C

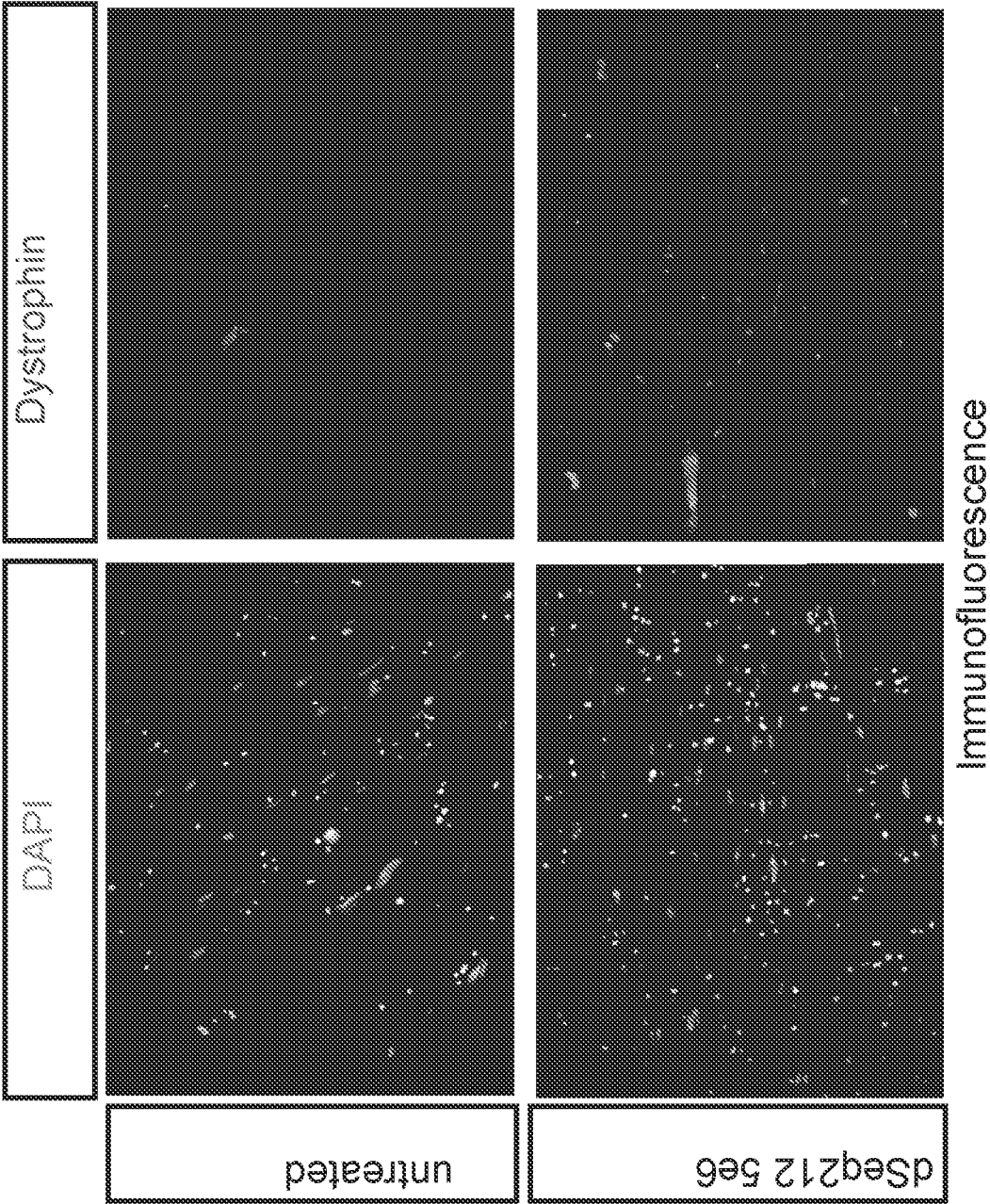


FIG. 26A

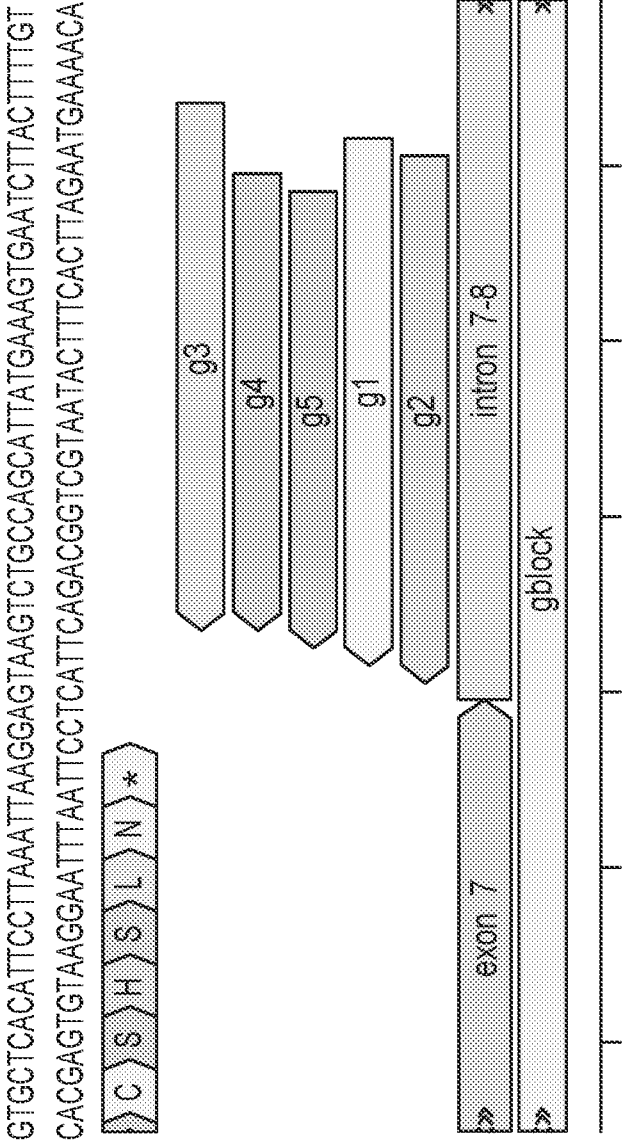


FIG. 26B

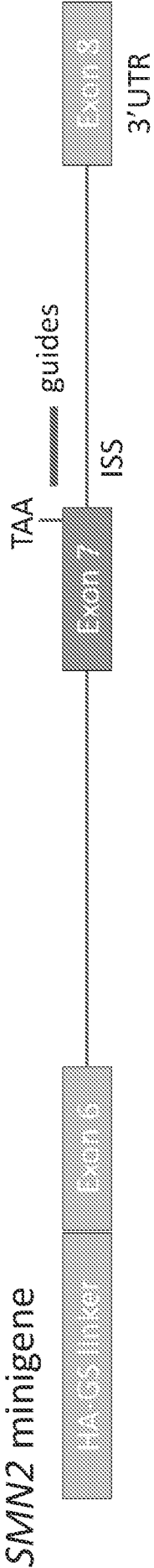


FIG. 26C

