

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property
Organization
International Bureau



(43) International Publication Date
8 June 2017 (08.06.2017)

(10) International Publication Number
WO 2017/095967 A2

(51) International Patent Classification:
C12Q 1/68 (2006.01)

(21) International Application Number:
PCT/US2016/064285

(22) International Filing Date:
30 November 2016 (30.11.2016)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/260,712 30 November 2015 (30.11.2015) US
62/330,336 2 May 2016 (02.05.2016) US

(71) Applicant: DUKE UNIVERSITY [US/US]; 2812 Erwin
Road, Suite 306, Durham, NC 27705 (US).

(72) Inventors; and

(71) Applicants : GERSBACH, Charles, A. [US/US]; 3220
Wilderness Road, Durham, NC 27712 (US). ROBINSON-
HAMM, Jacqueline, N. [US/US]; c/o Duke University,
2812 Erwin Road Suite 306, Durham, NC 27705 (US).

(74) Agent: YEH, Sansun; Michael Best & Friedrich LLP, 100
East Wisconsin Avenue, Suite 3300, Milwaukee, WI
53202-4108 (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, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IR, IS, JP, KE, KG, KN, KP, KR,
KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME,
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, SM, ST, SV, SY, TH, TJ, TM,
TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM,
ZW.

(84) Designated States (unless otherwise indicated, for every
kind of regional protection available): ARIPO (BW, 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, 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).

Published:

— with declaration under Article 17(2)(a); without abstract;
title not checked by the International Searching Authority

(54) Title: THERAPEUTIC TARGETS FOR THE CORRECTION OF THE HUMAN DYSTROPHIN GENE BY GENE EDITING
AND METHODS OF USE

(57) Abstract:



WO 2017/095967 A2

THERAPEUTIC TARGETS FOR THE CORRECTION OF THE HUMAN DYSTROPHIN GENE BY GENE EDITING AND METHODS OF USE

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application No. 62/260,712, filed November 30, 2015, and U.S. Provisional Application No. 62/330,336, filed May 2, 2016, all of which are incorporated herein by reference in their entirety.

STATEMENT OF GOVERNMENT INTEREST

[0002] This invention was made with government support under an award by the National Science Foundation Graduate Research Fellowship Program. The U.S. Government has certain rights to this invention.

TECHNICAL FIELD

[0003] The present disclosure relates to the field of gene expression alteration, genome engineering and genomic alteration of genes using Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-associated (Cas) 9-based systems and viral delivery systems. The present disclosure also relates to the field of genome engineering and genomic alteration of genes in muscle, such as skeletal muscle and cardiac muscle.

BACKGROUND

[0004] Synthetic transcription factors have been engineered to control gene expression for many different medical and scientific applications in mammalian systems, including stimulating tissue regeneration, drug screening, compensating for genetic defects, activating silenced tumor suppressors, controlling stem cell differentiation, performing genetic screens, and creating synthetic gene circuits. These transcription factors can target promoters or enhancers of endogenous genes, or be purposefully designed to recognize sequences orthogonal to mammalian genomes for transgene regulation. The most common strategies for engineering novel

transcription factors targeted to user-defined sequences have been based on the programmable DNA-binding domains of zinc finger proteins and transcription-activator like effectors (TALEs). Both of these approaches involve applying the principles of protein-DNA interactions of these domains to engineer new proteins with unique DNA-binding specificity. Although these methods have been widely successful for many applications, the protein engineering necessary for manipulating protein-DNA interactions can be laborious and require specialized expertise.

[0005] Additionally, these new proteins are not always effective. The reasons for this are not yet known but may be related to the effects of epigenetic modifications and chromatin state on protein binding to the genomic target site. In addition, there are challenges in ensuring that these new proteins, as well as other components, are delivered to each cell. Existing methods for delivering these new proteins and their multiple components include delivery to cells on separate plasmids or vectors which leads to highly variable expression levels in each cell due to differences in copy number. Additionally, gene activation following transfection is transient due to dilution of plasmid DNA, and temporary gene expression may not be sufficient for inducing therapeutic effects. Furthermore, this approach is not amenable to cell types that are not easily transfected. Thus another limitation of these new proteins is the potency of transcriptional activation.

[0006] CRISPR/Cas9-based gene editing systems can be used to introduce site-specific double strand breaks at targeted genomic loci. This DNA cleavage stimulates the natural DNA-repair machinery, leading to one of two possible repair pathways. In the absence of a donor template, the break will be repaired by non-homologous end joining (NHEJ), an error-prone repair pathway that leads to small insertions or deletions of DNA. This method can be used to intentionally disrupt, delete, or alter the reading frame of targeted gene sequences. However, if a donor template is provided along with the nucleases, then the cellular machinery will repair the break by homologous recombination, which is enhanced several orders of magnitude in the presence of DNA cleavage. This method can be used to introduce specific changes in the DNA sequence at target sites. Engineered nucleases have been used for gene editing in a variety of human stem cells and cell lines, and for gene editing in the mouse liver. However, the major hurdle for implementation of these technologies is delivery to particular tissues *in vivo* in a way that is effective, efficient, and facilitates successful genome modification.

[0007] Hereditary genetic diseases have devastating effects on children in the United States. These diseases currently have no cure and can only be managed by attempts to alleviate the symptoms. For decades, the field of gene therapy has promised a cure to these diseases. However technical hurdles regarding the safe and efficient delivery of therapeutic genes to cells and patients have limited this approach. Duchenne muscular dystrophy (DMD) is a fatal genetic disease, clinically characterized by muscle wasting, loss of ambulation, and death typically in the third decade of life due to the loss of functional dystrophin. DMD is the result of inherited or spontaneous mutations in the *dystrophin* gene. Most mutations causing DMD are a result of deletions of exon(s), pushing the translational reading frame out of frame.

[0008] Dystrophin is a key component of a protein complex that is responsible for regulating muscle cell integrity and function. DMD patients typically lose the ability to physically support themselves during childhood, become progressively weaker during the teenage years, and die in their twenties. Current experimental gene therapy strategies for DMD require repeated administration of transient gene delivery vehicles or rely on permanent integration of foreign genetic material into the genomic DNA. Both of these methods have serious safety concerns. Furthermore, these strategies have been limited by an inability to deliver the large and complex *dystrophin* gene sequence. There remains a need for more precise and efficient gene editing tools for correcting and treating patients with mutations in the *dystrophin* gene.

SUMMARY

[0009] The present invention is directed to a guide RNA (gRNA) comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42 or a complement thereof.

[0010] The present invention is also directed to a DNA targeting composition comprising a first gRNA and a second gRNA. The first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO:

13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42, or a complement thereof. The first gRNA molecule and the second gRNA molecule comprise different targeting domains.

[0011] The present invention is also directed to an isolated polynucleotide comprising the gRNA molecule described above or the DNA targeting composition described above.

[0012] The present invention is directed to a vector comprising the gRNA described above, the DNA targeting composition described above, or the isolated polynucleotide described above.

[0013] The present invention is also directed to a vector comprising the DNA targeting composition described above.

[0014] The present invention is also directed to a vector encoding: (a) a first guide RNA (gRNA) molecule, (b) a second gRNA molecule, and (c) at least one Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25). The first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42, or a complement thereof. The first gRNA molecule and the second gRNA molecule comprise different targeting domains.

[0015] The present invention is also directed to a cell comprising the gRNA described above, the DNA targeting composition described above, the isolated polynucleotide described above, or the vector of described above.

~~**[0016]**~~ The present invention is also directed to a kit comprising the gRNA described above, the DNA targeting system described above, the isolated polynucleotide described above, the vector described above, or the cell described above and optionally instructions for use.

[0017] The present invention is also directed to a method of correcting a mutant *dystrophin* gene in a cell. The method comprises administering to a cell the gRNA described above, the DNA targeting system described above, the isolated polynucleotide described above, or the vector described above.

[0018] The present invention is also directed to a method of genome editing a mutant *dystrophin* gene in a subject. The method comprises administering to the subject a genome

editing composition comprising the gRNA described above, the DNA targeting system described above, the isolated polynucleotide described above, the vector described above, or the cell described above.

[0019] The present invention is also directed to a method of treating a subject in need thereof having a mutant *dystrophin* gene. The method comprises administering to the subject the gRNA described above, the DNA targeting system described above, the isolated polynucleotide described above, the vector described above, or the cell described above.

[0020] The present invention is also directed to a modified adeno-associated viral vector for genome editing a mutant *dystrophin* gene in a subject comprising a first polynucleotide sequence encoding the gRNA described above, and a second polynucleotide sequence encoding a Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25).

[0021] The present invention is also directed to a composition for deleting a segment of a *dystrophin* gene comprising exon 51, the composition comprising: (a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and (b) a second vector comprising a polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25). Each of the first and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the first vector and second vector are configured to form a first and a second double strand break in a first intron and a second intron flanking exon 51 of the human *DMD* gene, respectively, thereby deleting a segment of the *dystrophin* gene comprising exon 51.

[0022] The present invention is also directed to a cell comprising the composition described above.

[0023] The present invention is also directed to a method of correcting a mutant *dystrophin* gene in a cell, comprising administering to the cell: (a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and (b) a second vector comprising a

polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25). Each of the first gRNA and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the vector is configured to form a first and a second double strand break in a first and a second intron flanking exon 51 of the human *dystrophin* gene, respectively, thereby deleting a segment of the *dystrophin* gene comprising exon 51.

[0024] The present invention is also directed to a method of treating a subject in need thereof having a mutant *dystrophin* gene. The method comprises administering to the subject: (a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and (b) a second vector comprising a polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25). Each of the first gRNA and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the vector is configured to form a first and a second double strand break in a first and a second intron flanking exon 51 of the human *dystrophin* gene, respectively, thereby deleting a segment of the *dystrophin* gene comprising exon 51.

[0025] The present invention is also directed to a method of generating a transgenic rodent embryo having a human *dystrophin* gene (hDMD) with an exon 52 deletion (Δ 52). The method comprises administering to a rodent embryo the gRNA described above, the DNA targeting system described above, the isolated polynucleotide described above, the vector described above, the modified adeno-associated viral vector described above, or the composition described above, thereby deleting exon 52 of the human *dystrophin* gene, and selecting for a transgenic rodent embryo having a deletion of exon 52 of the human *dystrophin* gene, wherein the rodent embryo comprises a normal human *dystrophin* gene.

[0026] The present invention is also directed to a transgenic rodent embryo produced by the method described above.

[0027] The present invention is also directed to a transgenic rodent produced from the transgenic rodent embryo described above.

BRIEF DESCRIPTION OF THE DRAWINGS

[0028] FIG. 1 shows the activity of individual gRNAs JCR89 and JCR91 that target the human *dystrophin* gene in HEK293T cells (wild-type *dystrophin* gene) and DMD patient myoblast lines (DMD 8036 and DMD 6594, each of which has a mutant form of the *dystrophin* gene), as determined by the Surveyor Assay.

[0029] FIGS. 2A and 2B show the deletion of exon 51 in genomic DNA of HEK293T cells and DMD myoblasts (DMD 8036 and DMD 6594) (FIG. 2A) and cDNA from DMD myoblasts myoblasts (FIG. 2B) by co-treatment with SaCas9 and gRNAs JCR89 and JCR91.

[0030] FIG. 3 shows an AAV-based *in vivo* system for co-delivery of SaCas9 and gRNAs JCR89 and JCR91 on two viral vectors to muscle tissues.

[0031] FIG. 4 shows the detection of the deletion of human exon 51 in transgenic mice carrying the human DMD gene (hDMD/*mdx* mice) following local AAV8 delivery of viral vectors carrying the SaCas9 and gRNAs to tibialis anterior (TA) muscle.

[0032] FIG. 5 shows the detection of the deletion of human exon 51 in transgenic mice carrying the hDMD gene following systemic AAV8 delivery via tail vein injection.

[0033] FIG. 6 shows the various gRNA targets that are conserved between human and rhesus macaque genomes (see sequences of gRNA in Table 2). The location of each gRNA is indicated in relation to Exon 51 of the human *dystrophin* gene.

[0034] FIG. 7 shows the activity of individual gRNAs following transfection of human HEK293T cells as determined by the Surveyor assay.

[0035] FIG. 8 shows the specificity of the candidate gRNAs as predicted using CasOFFinder program (Bae et al. (2014) *Bioinformatics* 30: 1473-1475).

[0036] FIG. 9 shows the deletion of exon 51 by gRNAs JCR157 and JCR160 in HEK293T cells and DMD 6594 cells as determined by PCR of genomic DNA.

[0037] FIG. 10 shows the activity, as determined by the Surveyor Assay, of various target lengths of gRNA JCR157: 19, 20, 21, 22, and 23 nucleotides.

[0038] FIG. 11 shows the activity, as determined by the Surveyor Assay, of various target lengths of gRNA JCR160: 19, 20, 21, 22, and 23 nucleotides.

[0039] FIG. 12 shows the deletions generated by combining JCR157 and JCR160 of various lengths (21, 22, or 23 nucleotides) as determined by PCR of genomic DNA.

[0040] FIG. 13 shows *in vitro* on-target nuclease activity by Surveyor assay.

- [0041] FIG. 14 shows the *in vitro* deletion of exon 51 in genomic DNA.
- [0042] FIG. 15 shows the *in vitro* deletion of exon 51 in cDNA in human DMD myoblasts differentiated for 7 days.
- [0043] FIG. 16 shows *in vitro* exon 47 to 52 junction in cDNA of DMD patient myoblasts.
- [0044] FIG. 17 shows the design for the $\Delta 52/mdx$ mouse starting from healthy hDMD/*mdx* mouse.
- [0045] FIG. 18 shows *in vitro* guide validation: individual (Surveyor assay).
- [0046] FIG. 19 shows *in vitro* guide validation: paired: the deletion of exon 51 in genomic DNA of HEK293T cells was generated using pairs of gRNAs.
- [0047] FIG. 20 shows schematic of DNA microinjection protocol.
- [0048] FIG. 21 shows a schematic of mouse breeding.
- [0049] FIG. 22 shows founder mice genotyping results.
- [0050] FIG. 23 shows a portion of sequencing results from founder mice 7, 63, and 76.
- [0051] FIG. 24 shows a schematic of further mouse breeding.
- [0052] FIG. 25 shows the genotyping of litter 5 (males only) from the founder male 76 + *mdx/mdx* breeding results. "63" is a founder male (but was not the parent in this case). "293" represents HEK293T cell genomic DNA control.
- [0053] FIG. 26 shows the genotyping of litter 1 from the founder male 63 + *mdx/mdx* breeding results.
- [0054] FIG. 27 shows a portion of a 392 bp sequencing read of pups 54497 and 54498.
- [0055] FIG. 28 shows immunohistochemical staining of heart and TA from pups 54497 and 54498.
- [0056] FIG. 29 shows that the $\Delta 52/mdx$ mouse lacks dystrophin protein.
- [0057] FIG. 30 shows a Western blot indicating that the $\Delta 52/mdx$ mouse lacks dystrophin protein which is consistent with the DMD genotype, while the healthy hDMD/*mdx* mouse expresses dystrophin.
- [0058] FIG. 31 shows the overall activity of the $\Delta 52/mdx$ mouse compared to *mdx* mice and hDMD/*mdx* mice as indicated by locomotion and exploration.
- [0059] FIG. 32 shows the corrections strategy for the $\Delta 52/mdx$ mouse using SaCas9 and gRNAs to skip exon 51 by targeting gRNAs upstream and downstream of exon 51 in the intronic region for removal.

[0060] FIG. 33 shows *in vitro* restoration of dystrophin protein from exon 51 deletion in DMD patient myoblasts (DMD 6594 cells) using SaCas9 and gRNAs JCR179 and JCR183.

[0061] FIG. 34 shows the experimental design to treat $\Delta 52/mdx$ mouse using the gRNAs and SaCas9 system.

[0062] FIG. 35 shows *in vivo* exon 51 deletion in right TA muscle.

[0063] FIG. 36 shows *in vivo* exon 51 deletion in right TA muscle.

[0064] FIG. 37 shows *in vivo* dystrophin protein restoration in treated TA muscle.

[0065] FIG. 38 shows *in vivo* dystrophin protein restoration in treated TA muscle.

[0066] FIG. 39 shows average of all time points for total distance moved.

[0067] FIG. 40 shows average of all time points for total rearing postures.

[0068] FIG. 41 shows the grip strength of 16 week untreated and treated mice.

[0069] FIG. 42 shows cDNA PCR results of heart tissue.

[0070] FIG. 43 shows sequencing of the amplified cDNA PCR bands from FIG. 42.

DETAILED DESCRIPTION

[0071] As described herein, certain methods and engineered gRNAs have been discovered to be useful with CRISPR/CRISPR-associated (Cas) 9-based gene editing systems for altering the expression, genome engineering, and correcting or reducing the effects of mutations in the *dystrophin* gene involved in genetic diseases, e.g., DMD. The disclosed gRNAs were generated to target sites that are more amenable to clinical translation. For example, the gene encoding *S. pyogenes* Cas9 (SpCas9) is too large to be delivered by adeno-associated virus (AAV), a vector used for the systemic gene delivery to muscle when all other necessary regulatory sequences are included. Instead, the disclosed gRNAs were selected and screened for use with *S. aureus* Cas9 (SaCas9), which is about 1 kb smaller than SpCas9. The target selections were screened for being SaCas9-compatible targets on sequences that were conserved between the human and rhesus macaque genomes, which greatly limits the number of possible gene targets. This selection criterion was chosen to allow for gRNA candidates that could be active in both humans and rhesus monkeys so as to facilitate preclinical testing in non-human primate models. The disclosed gRNAs, which target both human and rhesus monkey *dystrophin* gene sequences, can be used with the CRISPR/Cas9-based system to target intronic regions surrounding exon 51 of

the human *dystrophin* gene, causing genomic deletions of this region in order to restore expression of functional dystrophin in cells from DMD patients.

[0072] Also described herein are genetic constructs, compositions and methods for delivering CRISPR/Cas9-based gene editing system and multiple gRNAs to target the *dystrophin* gene. The presently disclosed subject matter also provides for methods for delivering the genetic constructs (e.g., vectors) or compositions comprising thereof to skeletal muscle and cardiac muscle. The vector can be an AAV, including modified AAV vectors. The presently disclosed subject matter describes a way to deliver active forms of this class of therapeutics to skeletal muscle or cardiac muscle that is effective, efficient and facilitates successful genome modification, as well as provide a means to rewrite the human genome for therapeutic applications and target model species for basic science applications.

[0073] Section headings as used in this section and the entire disclosure herein are merely for organizational purposes and are not intended to be limiting.

1. Definitions

[0074] 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. In case of conflict, the present document, including definitions, will control. Preferred methods and materials are described below, although methods and materials similar or equivalent to those described herein can be used in practice or testing of the present invention. All publications, patent applications, patents and other references mentioned herein are incorporated by reference in their entirety. The materials, methods, and examples disclosed herein are illustrative only and not intended to be limiting.

[0075] The terms “comprise(s),” “include(s),” “having,” “has,” “can,” “contain(s),” and variants thereof, as used herein, are intended to be open-ended transitional phrases, terms, or words that do not preclude the possibility of additional acts or structures. The singular forms “a,” “and” and “the” include plural references unless the context clearly dictates otherwise. The present disclosure also contemplates other embodiments “comprising,” “consisting of” and “consisting essentially of,” the embodiments or elements presented herein, whether explicitly set forth or not.

[0076] For the recitation of numeric ranges herein, each intervening number there between with the same degree of precision is explicitly contemplated. For example, for the range of 6-9,

the numbers 7 and 8 are contemplated in addition to 6 and 9, and for the range 6.0-7.0, the number 6.0, 6.1, 6.2, 6.3, 6.4, 6.5, 6.6, 6.7, 6.8, 6.9, and 7.0 are explicitly contemplated.

[0077] As used herein, the term “about” or “approximately” means within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, *i.e.*, the limitations of the measurement system. For example, “about” can mean within 3 or more than 3 standard deviations, per the practice in the art. Alternatively, “about” can mean a range of up to 20%, preferably up to 10%, more preferably up to 5%, and more preferably still up to 1% of a given value. Alternatively, particularly with respect to biological systems or processes, the term can mean within an order of magnitude, preferably within 5-fold, and more preferably within 2-fold, of a value.

[0078] “Adeno-associated virus” or “AAV” as used interchangeably herein refers to a small virus belonging to the genus Dependovirus of the Parvoviridae family that infects humans and some other primate species. AAV is not currently known to cause disease and consequently the virus causes a very mild immune response.

[0079] “Binding region” as used herein refers to the region within a nuclease target region that is recognized and bound by the nuclease.

[0080] “Cardiac muscle” or “heart muscle” as used interchangeably herein means a type of involuntary striated muscle found in the walls and histological foundation of the heart, the myocardium. Cardiac muscle is made of cardiomyocytes or myocardiocytes. Myocardiocytes show striations similar to those on skeletal muscle cells but contain only one, unique nucleus, unlike the multinucleated skeletal cells. In certain embodiments, “cardiac muscle condition” refers to a condition related to the cardiac muscle, such as cardiomyopathy, heart failure, arrhythmia, and inflammatory heart disease.

[0081] “Coding sequence” or “encoding nucleic acid” as used herein means the nucleic acids (RNA or DNA molecule) that comprise a nucleotide sequence which encodes a protein. The coding sequence can further include initiation and termination signals operably linked to regulatory elements including a promoter and polyadenylation signal capable of directing expression in the cells of an individual or mammal to which the nucleic acid is administered. The coding sequence may be codon optimize.

[0082] “Complement” or “complementary” as used herein means a nucleic acid can mean Watson-Crick (*e.g.*, A-T/U and C-G) or Hoogsteen base pairing between nucleotides or

nucleotide analogs of nucleic acid molecules. “Complementarity” refers to a property shared between two nucleic acid sequences, such that when they are aligned antiparallel to each other, the nucleotide bases at each position will be complementary.

[0083] “Correcting”, “genome editing” and “restoring” as used herein refers to changing a mutant gene that encodes a truncated protein or no protein at all, such that a full-length functional or partially full-length functional protein expression is obtained. Correcting or restoring a mutant gene may include replacing the region of the gene that has the mutation or replacing the entire mutant gene with a copy of the gene that does not have the mutation with a repair mechanism such as homology-directed repair (HDR). Correcting or restoring a mutant gene may also include repairing a frameshift mutation that causes a premature stop codon, an aberrant splice acceptor site or an aberrant splice donor site, by generating a double stranded break in the gene that is then repaired using non-homologous end joining (NHEJ). NHEJ may add or delete at least one base pair during repair which may restore the proper reading frame and eliminate the premature stop codon. Correcting or restoring a mutant gene may also include disrupting an aberrant splice acceptor site or splice donor sequence. Correcting or restoring a mutant gene may also include deleting a non-essential gene segment by the simultaneous action of two nucleases on the same DNA strand in order to restore the proper reading frame by removing the DNA between the two nuclease target sites and repairing the DNA break by NHEJ.

[0084] “Donor DNA”, “donor template” and “repair template” as used interchangeably herein refers to a double-stranded DNA fragment or molecule that includes at least a portion of the gene of interest. The donor DNA may encode a full-functional protein or a partially-functional protein.

[0085] “Duchenne Muscular Dystrophy” or “DMD” as used interchangeably herein refers to a recessive, fatal, X-linked disorder that results in muscle degeneration and eventual death. DMD is a common hereditary monogenic disease and occurs in 1 in 3500 males. DMD is the result of inherited or spontaneous mutations that cause nonsense or frame shift mutations in the *dystrophin* gene. The majority of dystrophin mutations that cause DMD are deletions of exons that disrupt the reading frame and cause premature translation termination in the *dystrophin* gene. DMD patients typically lose the ability to physically support themselves during childhood, become progressively weaker during the teenage years, and die in their twenties.

[0086] “Dystrophin” as used herein refers to a rod-shaped cytoplasmic protein which is a part of a protein complex that connects the cytoskeleton of a muscle fiber to the surrounding extracellular matrix through the cell membrane. Dystrophin provides structural stability to the dystroglycan complex of the cell membrane that is responsible for regulating muscle cell integrity and function. The *dystrophin* gene or “DMD gene” as used interchangeably herein is 2.2 megabases at locus Xp21. The primary transcription measures about 2,400 kb with the mature mRNA being about 14 kb. 79 exons code for the protein which is over 3500 amino acids.

[0087] “Exon 51” as used herein refers to the 51st exon of the *dystrophin* gene. Exon 51 is frequently adjacent to frame-disrupting deletions in DMD patients and has been targeted in clinical trials for oligonucleotide-based exon skipping. A clinical trial for the exon 51 skipping compound eteplirsen recently reported a significant functional benefit across 48 weeks, with an average of 47% dystrophin positive fibers compared to baseline. Mutations in exon 51 are ideally suited for permanent correction by NHEJ-based genome editing.

[0088] “Frameshift” or “frameshift mutation” as used interchangeably herein refers to a type of gene mutation wherein the addition or deletion of one or more nucleotides causes a shift in the reading frame of the codons in the mRNA. The shift in reading frame may lead to the alteration in the amino acid sequence at protein translation, such as a missense mutation or a premature stop codon.

[0089] “Functional” and “full-functional” as used herein describes protein that has biological activity. A “functional gene” refers to a gene transcribed to mRNA, which is translated to a functional protein.

[0090] “Fusion protein” as used herein refers to a chimeric protein created through the joining of two or more genes that originally coded for separate proteins. The translation of the fusion gene results in a single polypeptide with functional properties derived from each of the original proteins.

[0091] “Genetic construct” as used herein refers to the DNA or RNA molecules that comprise a nucleotide sequence that encodes a protein. The coding sequence includes initiation and termination signals operably linked to regulatory elements including a promoter and polyadenylation signal capable of directing expression in the cells of the individual to whom the nucleic acid molecule is administered. As used herein, the term “expressible form” refers to gene constructs that contain the necessary regulatory elements operable linked to a coding sequence

that encodes a protein such that when present in the cell of the individual, the coding sequence will be expressed.

[0092] “Genetic disease” as used herein refers to a disease, partially or completely, directly or indirectly, caused by one or more abnormalities in the genome, especially a condition that is present from birth. The abnormality may be a mutation, an insertion or a deletion. The abnormality may affect the coding sequence of the gene or its regulatory sequence. The genetic disease may be, but not limited to DMD, Becker Muscular Dystrophy (BMD), hemophilia, cystic fibrosis, Huntington's chorea, familial hypercholesterolemia (LDL receptor defect), hepatoblastoma, Wilson's disease, congenital hepatic porphyria, inherited disorders of hepatic metabolism, Lesch Nyhan syndrome, sickle cell anemia, thalassaemias, xeroderma pigmentosum, Fanconi's anemia, retinitis pigmentosa, ataxia telangiectasia, Bloom's syndrome, retinoblastoma, and Tay-Sachs disease.

[0093] “Homology-directed repair” or “HDR” as used interchangeably herein refers to a mechanism in cells to repair double strand DNA lesions when a homologous piece of DNA is present in the nucleus, mostly in G2 and S phase of the cell cycle. HDR uses a donor DNA template to guide repair and may be used to create specific sequence changes to the genome, including the targeted addition of whole genes. If a donor template is provided along with the CRISPR/Cas9-based gene editing system, then the cellular machinery will repair the break by homologous recombination, which is enhanced several orders of magnitude in the presence of DNA cleavage. When the homologous DNA piece is absent, non-homologous end joining may take place instead.

[0094] “Genome editing” as used herein refers to changing a gene. Genome editing may include correcting or restoring a mutant gene. Genome editing may include knocking out a gene, such as a mutant gene or a normal gene. Genome editing may be used to treat disease or enhance muscle repair by changing the gene of interest.

[0095] “Identical” or “identity” as used herein in the context of two or more nucleic acids or polypeptide sequences means that the sequences have a specified percentage of residues that are the same over a specified region. The percentage may be calculated by optimally aligning the two sequences, comparing the two sequences over the specified region, determining the number of positions at which the identical residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in

the specified region, and multiplying the result by 100 to yield the percentage of sequence identity. In cases where the two sequences are of different lengths or the alignment produces one or more staggered ends and the specified region of comparison includes only a single sequence, the residues of single sequence are included in the denominator but not the numerator of the calculation. When comparing DNA and RNA, thymine (T) and uracil (U) may be considered equivalent. Identity may be performed manually or by using a computer sequence algorithm such as BLAST or BLAST 2.0.

[0096] “Mutant gene” or “mutated gene” as used interchangeably herein refers to a gene that has undergone a detectable mutation. A mutant gene has undergone a change, such as the loss, gain, or exchange of genetic material, which affects the normal transmission and expression of the gene. A “disrupted gene” as used herein refers to a mutant gene that has a mutation that causes a premature stop codon. The disrupted gene product is truncated relative to a full-length undisrupted gene product.

[0097] “Non-homologous end joining (NHEJ) pathway” as used herein refers to a pathway that repairs double-strand breaks in DNA by directly ligating the break ends without the need for a homologous template. The template-independent re-ligation of DNA ends by NHEJ is a stochastic, error-prone repair process that introduces random micro-insertions and micro-deletions (indels) at the DNA breakpoint. This method may be used to intentionally disrupt, delete, or alter the reading frame of targeted gene sequences. NHEJ typically uses short homologous DNA sequences called microhomologies to guide repair. These microhomologies are often present in single-stranded overhangs on the end of double-strand breaks. When the overhangs are perfectly compatible, NHEJ usually repairs the break accurately, yet imprecise repair leading to loss of nucleotides may also occur, but is much more common when the overhangs are not compatible.

[0098] “Normal gene” as used herein refers to a gene that has not undergone a change, such as a loss, gain, or exchange of genetic material. The normal gene undergoes normal gene transmission and gene expression.

[0099] “Nuclease mediated NHEJ” as used herein refers to NHEJ that is initiated after a nuclease, such as a Cas9 molecule, cuts double stranded DNA.

[00100] “Nucleic acid” or “oligonucleotide” or “polynucleotide” as used herein means at least two nucleotides covalently linked together. The depiction of a single strand also defines the

sequence of the complementary strand. Thus, a nucleic acid also encompasses the complementary strand of a depicted single strand. Many variants of a nucleic acid may be used for the same purpose as a given nucleic acid. Thus, a nucleic acid also encompasses substantially identical nucleic acids and complements thereof. A single strand provides a probe that may hybridize to a target sequence under stringent hybridization conditions. Thus, a nucleic acid also encompasses a probe that hybridizes under stringent hybridization conditions.

[00101] Nucleic acids may be single stranded or double stranded, or may contain portions of both double stranded and single stranded sequence. The nucleic acid may be DNA, both genomic and cDNA, RNA, or a hybrid, where the nucleic acid may contain combinations of deoxyribo- and ribo-nucleotides, and combinations of bases including uracil, adenine, thymine, cytosine, guanine, inosine, xanthine hypoxanthine, isocytosine and isoguanine. Nucleic acids may be obtained by chemical synthesis methods or by recombinant methods.

[00102] “Operably linked” as used herein means that expression of a gene is under the control of a promoter with which it is spatially connected. A promoter may be positioned 5' (upstream) or 3' (downstream) of a gene under its control. The distance between the promoter and a gene may be approximately the same as the distance between that promoter and the gene it controls in the gene from which the promoter is derived. As is known in the art, variation in this distance may be accommodated without loss of promoter function.

[00103] “Partially-functional” as used herein describes a protein that is encoded by a mutant gene and has less biological activity than a functional protein but more than a non-functional protein.

[00104] “Premature stop codon” or “out-of-frame stop codon” as used interchangeably herein refers to nonsense mutation in a sequence of DNA, which results in a stop codon at location not normally found in the wild-type gene. A premature stop codon may cause a protein to be truncated or shorter compared to the full-length version of the protein.

[00105] “Promoter” as used herein means a synthetic or naturally-derived molecule which is capable of conferring, activating or enhancing expression of a nucleic acid in a cell. A promoter may comprise one or more specific transcriptional regulatory sequences to further enhance expression and/or to alter the spatial expression and/or temporal expression of same. A promoter may also comprise distal enhancer or repressor elements, which may be located as much as several thousand base pairs from the start site of transcription. A promoter may be derived from

sources including viral, bacterial, fungal, plants, insects, and animals. A promoter may regulate the expression of a gene component constitutively, or differentially with respect to cell, the tissue or organ in which expression occurs or, with respect to the developmental stage at which expression occurs, or in response to external stimuli such as physiological stresses, pathogens, metal ions, or inducing agents. Representative examples of promoters include the bacteriophage T7 promoter, bacteriophage T3 promoter, SP6 promoter, lac operator-promoter, tac promoter, SV40 late promoter, SV40 early promoter, RSV-LTR promoter, CMV IE promoter, SV40 early promoter or SV40 late promoter, human U6 (hU6) promoter, and CMV IE promoter.

[00106] “Skeletal muscle” as used herein refers to a type of striated muscle, which is under the control of the somatic nervous system and attached to bones by bundles of collagen fibers known as tendons. Skeletal muscle is made up of individual components known as myocytes, or “muscle cells”, sometimes colloquially called “muscle fibers.” Myocytes are formed from the fusion of developmental myoblasts (a type of embryonic progenitor cell that gives rise to a muscle cell) in a process known as myogenesis. These long, cylindrical, multinucleated cells are also called myofibers.

[00107] “Skeletal muscle condition” as used herein refers to a condition related to the skeletal muscle, such as muscular dystrophies, aging, muscle degeneration, wound healing, and muscle weakness or atrophy.

[00108] “Subject” and “patient” as used herein interchangeably refers to any vertebrate, including, but not limited to, a mammal (*e.g.*, cow, pig, camel, llama, horse, goat, rabbit, sheep, hamsters, guinea pig, cat, dog, rat, and mouse, a non-human primate (for example, a monkey, such as a cynomolgous or rhesus monkey, chimpanzee, etc.) and a human). In some embodiments, the subject may be a human or a non-human. The subject or patient may be undergoing other forms of treatment.

[00109] “Target gene” as used herein refers to any nucleotide sequence encoding a known or putative gene product. The target gene may be a mutated gene involved in a genetic disease. In certain embodiments, the target gene is a human *dystrophin* gene. In certain embodiments, the target gene is a mutant human *dystrophin* gene.

[00110] “Target region” as used herein refers to the region of the target gene to which the CRISPR/Cas9-based gene editing system is designed to bind and cleave.

[00111] “Transgene” as used herein refers to a gene or genetic material containing a gene sequence that has been isolated from one organism and is introduced into a different organism. This non-native segment of DNA may retain the ability to produce RNA or protein in the transgenic organism, or it may alter the normal function of the transgenic organism's genetic code. The introduction of a transgene has the potential to change the phenotype of an organism.

[00112] “Variant” used herein with respect to a nucleic acid means (i) a portion or fragment of a referenced nucleotide sequence; (ii) the complement of a referenced nucleotide sequence or portion thereof; (iii) a nucleic acid that is substantially identical to a referenced nucleic acid or the complement thereof; or (iv) a nucleic acid that hybridizes under stringent conditions to the referenced nucleic acid, complement thereof, or a sequences substantially identical thereto.

[00113] “Variant” with respect to a peptide or polypeptide that differs in amino acid sequence by the insertion, deletion, or conservative substitution of amino acids, but retain at least one biological activity. Variant may also mean a protein with an amino acid sequence that is substantially identical to a referenced protein with an amino acid sequence that retains at least one biological activity. A conservative substitution of an amino acid, *i.e.*, replacing an amino acid with a different amino acid of similar properties (*e.g.*, hydrophilicity, degree and distribution of charged regions) is recognized in the art as typically involving a minor change. These minor changes may be identified, in part, by considering the hydropathic index of amino acids, as understood in the art. Kyte *et al.*, *J. Mol. Biol.* 157:105-132 (1982). The hydropathic index of an amino acid is based on a consideration of its hydrophobicity and charge. It is known in the art that amino acids of similar hydropathic indexes may be substituted and still retain protein function. In one aspect, amino acids having hydropathic indexes of ± 2 are substituted. The hydrophilicity of amino acids may also be used to reveal substitutions that would result in proteins retaining biological function. A consideration of the hydrophilicity of amino acids in the context of a peptide permits calculation of the greatest local average hydrophilicity of that peptide. Substitutions may be performed with amino acids having hydrophilicity values within ± 2 of each other. Both the hydrophobicity index and the hydrophilicity value of amino acids are influenced by the particular side chain of that amino acid. Consistent with that observation, amino acid substitutions that are compatible with biological function are understood to depend on the relative similarity of the amino acids, and particularly the side chains of those amino acids, as revealed by the hydrophobicity, hydrophilicity, charge, size, and other properties.

[00114] "Vector" as used herein means a nucleic acid sequence containing an origin of replication. A vector may be a viral vector, bacteriophage, bacterial artificial chromosome or yeast artificial chromosome. A vector may be a DNA or RNA vector. A vector may be a self-replicating extrachromosomal vector, and preferably, is a DNA plasmid. For example, the vector may encode a Cas9 protein and at least one gRNA molecule, such as a gRNA comprising a targeting domain of any one of SEQ ID NOs: 1-19, 41, 42, or complement thereof. In some embodiments, the Cas9 protein may have an amino acid sequence of SEQ ID NO: 27, SEQ ID NO: 33, or SEQ ID NO: 45. In some embodiments, the Cas9 protein may be a *S. aureus* Cas9, such as a SaCas9 having an amino acid sequence of SEQ ID NO: 33 or 45. In some embodiments, the Cas9 protein is encoded by a nucleic acid sequence comprising a nucleic sequence of SEQ ID NO: 26, SEQ ID NO: 28, SEQ ID NO: 29, SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 32, SEQ ID NO: 34, SEQ ID NO: 35, SEQ ID NO: 43, or SEQ ID NO: 44.

[00115] Unless otherwise defined herein, scientific and technical terms used in connection with the present disclosure shall have the meanings that are commonly understood by those of ordinary skill in the art. For example, any nomenclatures used in connection with, and techniques of, cell and tissue culture, molecular biology, immunology, microbiology, genetics and protein and nucleic acid chemistry and hybridization described herein are those that are well known and commonly used in the art. The meaning and scope of the terms should be clear; in the event however of any latent ambiguity, definitions provided herein take precedent over any dictionary or extrinsic definition. Further, unless otherwise required by context, singular terms shall include pluralities and plural terms shall include the singular.

2. Genetic Constructs for Genome Editing of *Dystrophin* Gene

[00116] The present invention is directed to genetic constructs for genome editing, genomic alteration or altering gene expression of a *dystrophin* gene (e.g., human *dystrophin* gene). The genetic constructs include at least one gRNA that targets both human and rhesus monkey *dystrophin* gene sequences, such as SaCas9-compatible targets. The disclosed gRNAs can be included in a CRISPR/Cas9-based gene editing system, including systems that use SaCas9, to target intronic regions surrounding exon 51 of the human *dystrophin* gene, causing genomic deletions of this region in order to restore expression of functional dystrophin in cells from DMD patients.

a. *Dystrophin* Gene

[00117] Dystrophin is a rod-shaped cytoplasmic protein which is a part of a protein complex that connects the cytoskeleton of a muscle fiber to the surrounding extracellular matrix through the cell membrane. Dystrophin provides structural stability to the dystroglycan complex of the cell membrane. The *dystrophin* gene is 2.2 megabases at locus Xp21. The primary transcription measures about 2,400 kb with the mature mRNA being about 14 kb. 79 exons code for the protein which is over 3500 amino acids. Normal skeleton muscle tissue contains only small amounts of dystrophin but its absence or abnormal expression leads to the development of severe and incurable symptoms. Some mutations in the *dystrophin* gene lead to the production of defective dystrophin and severe dystrophic phenotype in affected patients. Some mutations in the *dystrophin* gene lead to partially-functional dystrophin protein and a much milder dystrophic phenotype in affected patients.

[00118] DMD is the result of inherited or spontaneous mutations that cause nonsense or frame shift mutations in the *dystrophin* gene. Naturally occurring mutations and their consequences are relatively well understood for DMD. It is known that in-frame deletions that occur in the exon 45-55 regions (e.g., exon 51) contained within the rod domain can produce highly functional dystrophin proteins, and many carriers are asymptomatic or display mild symptoms. Furthermore, more than 60% of patients may theoretically be treated by targeting exons in this region of the *dystrophin* gene (e.g., targeting exon 51). Efforts have been made to restore the disrupted dystrophin reading frame in DMD patients by skipping non-essential exon(s) (e.g., exon 51 skipping) during mRNA splicing to produce internally deleted but functional dystrophin proteins. The deletion of internal dystrophin exon(s) (e.g., deletion of exon 51) retains the proper reading frame but cause the less severe Becker muscular dystrophy, or BMD. The Becker muscular dystrophy, or BMD, genotype is similar to DMD in that deletions are present in the *dystrophin* gene. However, these deletions leave the reading frame intact. Thus an internally truncated but partially functional dystrophin protein is created. BMD has a wide array of phenotypes, but often if deletions are between exons 45-55 of dystrophin the phenotype is much milder compared to DMD. Thus changing a DMD genotype to a BMD genotype is a common strategy to correct dystrophin. There are many strategies to correct dystrophin, many of which rely on restoring the reading frame of the endogenous dystrophin. This shifts the disease genotype from DMD to Becker muscular dystrophy. Many BMD patients have intragenic

deletions that maintain the translational reading frame, leading to a shorter but largely functional dystrophin protein.

[00119] In certain embodiments, modification of exon 51 (e.g., deletion or excision of exon 51 by, e.g., NHEJ) to restore reading frame ameliorates the phenotype DMD subjects, including DMD subjects with deletion mutations. In certain embodiments, exon 51 of a *dystrophin* gene refers to the 51st exon of the *dystrophin* gene. Exon 51 is frequently adjacent to frame-disrupting deletions in DMD patients and has been targeted in clinical trials for oligonucleotide-based exon skipping. A clinical trial for the exon 51 skipping compound eteplirsen reported a significant functional benefit across 48 weeks, with an average of 47% dystrophin positive fibers compared to baseline. Mutations in exon 51 are ideally suited for permanent correction by NHEJ-based genome editing.

[00120] The presently disclosed vectors can generate deletions in the *dystrophin* gene, e.g., the human *dystrophin* gene. In certain embodiments, the vector is configured to form two double strand breaks (a first double strand break and a second double strand break) in two introns (a first intron and a second intron) flanking a target position of the *dystrophin* gene, thereby deleting a segment of the *dystrophin* gene comprising the dystrophin target position. A “dystrophin target position” can be a dystrophin exonic target position or a dystrophin intra-exonic target position, as described herein. Deletion of the *dystrophin* exonic target position can optimize the *dystrophin* sequence of a subject suffering from Duchenne muscular dystrophy, e.g., it can increase the function or activity of the encoded dystrophin protein, or results in an improvement in the disease state of the subject. In certain embodiments, excision of the *dystrophin* exonic target position restores reading frame. The *dystrophin* exonic target position can comprise one or more exons of the *dystrophin* gene. In certain embodiments, the dystrophin target position comprises exon 51 of the *dystrophin* gene (e.g., human *dystrophin* gene).

[00121] A presently disclosed genetic construct (e.g., a vector) can mediate highly efficient gene editing at exon 51 of a *dystrophin* gene (e.g., the human *dystrophin* gene). A presently disclosed genetic construct (e.g., a vector) restores dystrophin protein expression in cells from DMD patients.

[00122] Exon 51 is frequently adjacent to frame-disrupting deletions in DMD. Elimination of exon 51 from the *dystrophin* transcript by exon skipping can be used to treat approximately 15% of all DMD patients. This class of *dystrophin* mutations is ideally suited for permanent

correction by NHEJ-based genome editing and HDR. The genetic constructs (e.g., vectors) described herein have been developed for targeted modification of exon 51 in the human *dystrophin* gene. A presently disclosed genetic construct (e.g., a vector) is transfected into human DMD cells and mediates efficient gene modification and conversion to the correct reading frame. Protein restoration is concomitant with frame restoration and detected in a bulk population of CRISPR/Cas9-based gene editing system-treated cells.

b. CRISPR system

[00123] A presently disclosed genetic construct (e.g., a vector) encodes a CRISPR/Cas9-based gene editing system that is specific for a *dystrophin* gene (e.g., human *dystrophin* gene).

“Clustered Regularly Interspaced Short Palindromic Repeats” and “CRISPRs”, as used interchangeably herein refers to loci containing multiple short direct repeats that are found in the genomes of approximately 40% of sequenced bacteria and 90% of sequenced archaea. The CRISPR system is a microbial nuclease system involved in defense against invading phages and plasmids that provides a form of acquired immunity. The CRISPR loci in microbial hosts contain a combination of CRISPR-associated (Cas) genes as well as non-coding RNA elements capable of programming the specificity of the CRISPR-mediated nucleic acid cleavage. Short segments of foreign DNA, called spacers, are incorporated into the genome between CRISPR repeats, and serve as a 'memory' of past exposures. Cas9 forms a complex with the 3' end of the sgRNA (also referred interchangeably herein as “gRNA”), and the protein-RNA pair recognizes its genomic target by complementary base pairing between the 5' end of the sgRNA sequence and a predefined 20 bp DNA sequence, known as the protospacer. This complex is directed to homologous loci of pathogen DNA via regions encoded within the crRNA, i.e., the protospacers, and protospacer-adjacent motifs (PAMs) within the pathogen genome. The non-coding CRISPR array is transcribed and cleaved within direct repeats into short crRNAs containing individual spacer sequences, which direct Cas nucleases to the target site (protospacer). By simply exchanging the 20 bp recognition sequence of the expressed sgRNA, the Cas9 nuclease can be directed to new genomic targets. CRISPR spacers are used to recognize and silence exogenous genetic elements in a manner analogous to RNAi in eukaryotic organisms.

[00124] Three classes of CRISPR systems (Types I, II and III effector systems) are known. The Type II effector system carries out targeted DNA double-strand break in four sequential steps, using a single effector enzyme, Cas9, to cleave dsDNA. Compared to the Type I and Type

III effector systems, which require multiple distinct effectors acting as a complex, the Type II effector system may function in alternative contexts such as eukaryotic cells. The Type II effector system consists of a long pre-crRNA, which is transcribed from the spacer-containing CRISPR locus, the Cas9 protein, and a tracrRNA, which is involved in pre-crRNA processing. The tracrRNAs hybridize to the repeat regions separating the spacers of the pre-crRNA, thus initiating dsRNA cleavage by endogenous RNase III. This cleavage is followed by a second cleavage event within each spacer by Cas9, producing mature crRNAs that remain associated with the tracrRNA and Cas9, forming a Cas9:crRNA-tracrRNA complex.

[00125] The Cas9:crRNA-tracrRNA complex unwinds the DNA duplex and searches for sequences matching the crRNA to cleave. Target recognition occurs upon detection of complementarity between a “protospacer” sequence in the target DNA and the remaining spacer sequence in the crRNA. Cas9 mediates cleavage of target DNA if a correct protospacer-adjacent motif (PAM) is also present at the 3’ end of the protospacer. For protospacer targeting, the sequence must be immediately followed by the protospacer-adjacent motif (PAM), a short sequence recognized by the Cas9 nuclease that is required for DNA cleavage. Different Type II systems have differing PAM requirements. The *S. pyogenes* CRISPR system may have the PAM sequence for this Cas9 (SpCas9) as 5’-NRG-3’, where R is either A or G, and characterized the specificity of this system in human cells. A unique capability of the CRISPR/Cas9-based gene editing system is the straightforward ability to simultaneously target multiple distinct genomic loci by co-expressing a single Cas9 protein with two or more sgRNAs. For example, the *Streptococcus pyogenes* Type II system naturally prefers to use an “NGG” sequence, where “N” can be any nucleotide, but also accepts other PAM sequences, such as “NAG” in engineered systems (Hsu et al., *Nature Biotechnology* (2013) doi:10.1038/nbt.2647). Similarly, the Cas9 derived from *Neisseria meningitidis* (NmCas9) normally has a native PAM of NNNNGATT, but has activity across a variety of PAMs, including a highly degenerate NNNNGNNN PAM (Esvelt et al. *Nature Methods* (2013) doi:10.1038/nmeth.2681).

[00126] A Cas9 molecule of *S. aureus* recognizes the sequence motif NNGRR (R = A or G) (SEQ ID NO: 22) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence. In certain embodiments, a Cas9 molecule of *S. aureus* recognizes the sequence motif NNGRRN (R = A or G) (SEQ ID NO: 23) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence. In certain

embodiments, a Cas9 molecule of *S. aureus* recognizes the sequence motif NNGRRT (R = A or G) (SEQ ID NO: 24) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence. In certain embodiments, a Cas9 molecule of *S. aureus* recognizes the sequence motif NNGRRV (R = A or G) (SEQ ID NO: 25) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence. In the aforementioned embodiments, N can be any nucleotide residue, e.g., any of A, G, C, or T. Cas9 molecules can be engineered to alter the PAM specificity of the Cas9 molecule.

(1) CRISPR/Cas9-based gene editing system

[00127] An engineered form of the Type II effector system of *Streptococcus pyogenes* was shown to function in human cells for genome engineering. In this system, the Cas9 protein was directed to genomic target sites by a synthetically reconstituted “guide RNA” (“gRNA”, also used interchangeably herein as a chimeric single guide RNA (“sgRNA”)), which is a crRNA-tracrRNA fusion that obviates the need for RNase III and crRNA processing in general. Provided herein are CRISPR/Cas9-based engineered systems for use in genome editing and treating genetic diseases. The CRISPR/Cas9-based engineered systems can be designed to target any gene, including genes involved in a genetic disease, aging, tissue regeneration, or wound healing. The CRISPR/Cas9-based gene editing systems can include a Cas9 protein or Cas9 fusion protein and at least one gRNA. In certain embodiments, the system comprises two gRNA molecules. The Cas9 fusion protein may, for example, include a domain that has a different activity than what is endogenous to Cas9, such as a transactivation domain.

[00128] The target gene (e.g., a *dystrophin* gene, e.g., human *dystrophin* gene) can be involved in differentiation of a cell or any other process in which activation of a gene can be desired, or can have a mutation such as a frameshift mutation or a nonsense mutation. If the target gene has a mutation that causes a premature stop codon, an aberrant splice acceptor site or an aberrant splice donor site, the CRISPR/Cas9-based gene editing system can be designed to recognize and bind a nucleotide sequence upstream or downstream from the premature stop codon, the aberrant splice acceptor site or the aberrant splice donor site. The CRISPR-Cas9-based system can also be used to disrupt normal gene splicing by targeting splice acceptors and donors to induce skipping of premature stop codons or restore a disrupted reading frame. The CRISPR/Cas9-based gene editing system may or may not mediate off-target changes to protein-coding regions of the genome.

(a) Cas9 molecules and Cas9 fusion proteins

[00129] The CRISPR/Cas9-based gene editing system can include a Cas9 protein or a Cas9 fusion protein. Cas9 protein is an endonuclease that cleaves nucleic acid and is encoded by the CRISPR loci and is involved in the Type II CRISPR system. The Cas9 protein can be from any bacterial or archaea species, including, but not limited to, *Streptococcus pyogenes*, *Staphylococcus aureus* (*S. aureus*), *Acidovorax avenae*, *Actinobacillus pleuropneumoniae*, *Actinobacillus succinogenes*, *Actinobacillus suis*, *Actinomyces* sp., *Cyclophilus denitrificans*, *Aminomonas paucivorans*, *Bacillus cereus*, *Bacillus smithii*, *Bacillus thuringiensis*, *Bacteroides* sp., *Blastopirellula marina*, *Bradyrhizobium* sp., *Brevibacillus laterosporus*, *Campylobacter coli*, *Campylobacter jejuni*, *Campylobacter lari*, *Candidatus Puniceispirillum*, *Clostridium cellulolyticum*, *Clostridium perfringens*, *Corynebacterium accolens*, *Corynebacterium diphtheria*, *Corynebacterium matruchotii*, *Dinoroseobacter shibae*, *Eubacterium dolichum*, *gamma proteobacterium*, *Gluconacetobacter diazotrophicus*, *Haemophilus parainfluenzae*, *Haemophilus sputorum*, *Helicobacter canadensis*, *Helicobacter cinaedi*, *Helicobacter mustelae*, *Ilyobacter polytropus*, *Kingella kingae*, *Lactobacillus crispatus*, *Listeria ivanovii*, *Listeria monocytogenes*, *Listeriaceae bacterium*, *Methylocystis* sp., *Methylosinus trichosporium*, *Mobiluncus mulieris*, *Neisseria bacilliformis*, *Neisseria cinerea*, *Neisseria flavescens*, *Neisseria lactamica*, *Neisseria* sp., *Neisseria wadsworthii*, *Nitrosomonas* sp., *Parvibaculum lavamentivorans*, *Pasteurella multocida*, *Phascolarctobacterium succinatutens*, *Ralstonia syzygii*, *Rhodopseudomonas palustris*, *Rhodovulum* sp., *Simonsiella muelleri*, *Sphingomonas* sp., *Sporolactobacillus vineae*, *Staphylococcus lugdunensis*, *Streptococcus* sp., *Subdoligranulum* sp., *Tistrella mobilis*, *Treponema* sp., or *Verminephrobacter eiseniae*. In certain embodiments, the Cas9 molecule is a The Cas9 protein is a *Streptococcus pyogenes* Cas9 molecule (also referred herein as “SpCas9”). In certain embodiments, the Cas9 molecule is a *Staphylococcus aureus* Cas9 molecule (also referred herein as “SaCas9”).

[00130] A Cas9 molecule or a Cas9 fusion protein can interact with one or more gRNA molecule and, in concert with the gRNA molecule(s), localizes to a site which comprises a target domain, and in certain embodiments, a PAM sequence. The ability of a Cas9 molecule or a Cas9 fusion protein to recognize a PAM sequence can be determined, e.g., using a transformation assay as described previously (Jinek 2012).

[00131] In certain embodiments, the ability of a Cas9 molecule or a Cas9 fusion protein to interact with and cleave a target nucleic acid is PAM sequence dependent. A PAM sequence is a sequence in the target nucleic acid. In certain embodiments, cleavage of the target nucleic acid occurs upstream from the PAM sequence. Cas9 molecules from different bacterial species can recognize different sequence motifs (e.g., PAM sequences). In certain embodiments, a Cas9 molecule of *S. pyogenes* recognizes the sequence motif NGG and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence (see, e.g., Mali 2013). In certain embodiments, a Cas9 molecule of *S. thermophilus* recognizes the sequence motif NGGNG (SEQ ID NO: 36) and/or NNAGAAW (W = A or T) (SEQ ID NO: 20) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from these sequences (see, e.g., Horvath 2010; Deveau 2008). In certain embodiments, a Cas9 molecule of *S. mutans* recognizes the sequence motif NGG and/or NAAR (R = A or G) (SEQ ID NO: 21) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5 bp, upstream from this sequence (see, e.g., Deveau 2008). In certain embodiments, a Cas9 molecule of *S. aureus* recognizes the sequence motif NNGRR (R = A or G) (SEQ ID NO: 22) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence. In certain embodiments, a Cas9 molecule of *S. aureus* recognizes the sequence motif NNGRRN (R = A or G) (SEQ ID NO: 23) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence. In certain embodiments, a Cas9 molecule of *S. aureus* recognizes the sequence motif NNGRRT (R = A or G) (SEQ ID NO: 24) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence. In certain embodiments, a Cas9 molecule of *S. aureus* recognizes the sequence motif NNGRRV (R = A or G; V = A or C or G) (SEQ ID NO: 25) and directs cleavage of a target nucleic acid sequence 1 to 10, e.g., 3 to 5, bp upstream from that sequence. In the aforementioned embodiments, N can be any nucleotide residue, e.g., any of A, G, C, or T. Cas9 molecules can be engineered to alter the PAM specificity of the Cas9 molecule.

[00132] In certain embodiments, the vector encodes at least one Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25). In certain embodiments, the at least one Cas9 molecule is an *S. aureus* Cas9 molecule. In certain embodiments, the at least one Cas9 molecule is a mutant *S. aureus* Cas9 molecule.

[00133] The Cas9 protein can be mutated so that the nuclease activity is inactivated. An inactivated Cas9 protein (“iCas9”, also referred to as “dCas9”) with no endonuclease activity has been recently targeted to genes in bacteria, yeast, and human cells by gRNAs to silence gene expression through steric hindrance. Exemplary mutations with reference to the *S. pyogenes* Cas9 sequence include: D10A, E762A, H840A, N854A, N863A and/or D986A. Exemplary mutations with reference to the *S. aureus* Cas9 sequence include D10A and N580A. In certain embodiments, the Cas9 molecule is a mutant *S. aureus* Cas9 molecule. In certain embodiments, the mutant *S. aureus* Cas9 molecule comprises a D10A mutation. The nucleotide sequence encoding this mutant *S. aureus* Cas9 is set forth in SEQ ID NO: 34, which is provided below:

```

atgaaaagga actacattct ggggctggcc atcgggatta caagcgtggg gtatgggatt
attgactatg aaacaaggga cgtgatcgac gcaggcgtca gactgttcaa ggaggccaac
gtggaaaaca atgagggacg gagaagcaag aggggagcca ggcgcctgaa acgacggaga
aggcacagaa tccagagggg gaagaaactg ctgttcgatt acaacctgct gaccgaccat
tctgagctga gtggaattaa tccttatgaa gccagggtga aaggcctgag tcagaagctg
tcagaggaag agttttccgc agctctgctg cacctggcta agcgccgagg agtgcataac
gtcaatgagg tggaagagga caccggcaac gagctgtcta caaaggaaca gatctcacgc
aatagcaaag ctctggaaga gaagtatgtc gcagagctgc agctggaacg gctgaagaaa
gatggcgagg tgagaggggtc aattaatagg ttcaagacaa gcgactacgt caaagaagcc
aagcagctgc tgaaagtgca gaaggcttac caccagctgg atcagagctt catcgatact
tatatcgacc tgctggagac tcggagaacc tactatgagg gaccaggaga agggagcccc
ttcggatgga aagacatcaa ggaatggtag gagatgctga tgggacattg cacctatattt
ccagaagagc tgagaagcgt caagtacgct tataacgcag atctgtacaa cgccctgaat
gacctgaaca acctgggtcat caccagggat gaaaacgaga aactggaata ctatgagaag
ttccagatca tcgaaaacgt gtttaagcag aagaaaaagc ctacactgaa acagattgct
aaggagatcc tgggtcaacga agaggacatc aagggtacc ggggtgacaag cactggaaaa
ccagagttca ccaatctgaa agtgtatcac gatattaagg acatcacagc acggaaagaa
atcattgaga acgccgaact gctggatcag attgctaaga tcctgactat ctaccagagc
tccgaggaca tccaggaaga gctgactaac ctgaacagcg agctgacca ggaagagatc
gaacagatta gtaatctgaa ggggtacacc ggaacacaca acctgtccct gaaagctatc
aatctgattc tggatgagct gtggcataca aacgacaatc agattgcaat ctttaaccgg
ctgaagctgg tccccaaaaa ggtggacctg agtcagcaga aagagatccc aaccacactg
gtggacgatt tcattctgtc acccgtgggtc aagcggagct tcattccagag catcaaagtg
atcaacgcca tcattcaagaa gtacggcctg cccaatgata tcattatcga gctggctagg
gagaagaaca gcaaggacgc acagaagatg atcaatgaga tgcagaaacg aaaccggcag

```


accaatgaac gcattgaaga gattatccga actaccggga aagagaacgc aaagtacctg
 attgaaaaaa tcaagctgca cgatatgcag gagggaaagt gtctgtattc tctggaggcc
 atccccctgg aggacctgct gaacaatcca ttcaactacg aggtcgatca tattatcccc
 agaagcgtgt ccttcgacaa ttcctttaac aacaagggtgc tgggtcaagca ggaagagaac
 tctaaaaagg gcaataggac tcctttccag tacctgtcta gttcagattc caagatctct
 tacgaaacct ttaaaaagca cattctgaat ctggccaaag gaaagggccg catcagcaag
 accaaaaagg agtacctgct ggaagagcgg gacatcaaca gattctccgt ccagaaggat
 tttattaacc ggaatctggt ggacacaaga tacgctactc gcggcctgat gaatctgctg
 cgatcctatt tccgggtgaa caatctggat gtgaaagtca agtccatcaa cggcgggttc
 acatcttttc tgaggcgcaa atggaagttt aaaaaggagc gcaacaaagg gtacaagcac
 catgccgaag atgctctgat tatcgcaaat gccgacttca tctttaagga gtggaaaaag
 ctggacaaaag ccaagaaagt gatggagaac cagatgttcg aagagaagca ggccgaatct
 atgcccgaag tcgagacaga acaggagtac aaggagattt tcatcactcc tcaccagatc
 aagcatatca aggatttcaa ggactacaag tactctcacc ggggtggataa aaagcccaac
 agagagctga tcaatgacac cctgtatagt acaagaaaag acgataaggg gaataccctg
 attgtgaaca atctgaacgg actgtacgac aaagataatg acaagctgaa aaagctgac
 aacaaaagtc ccgagaagct gctgatgtac caccatgatc ctacagacata tcagaaactg
 aagctgatta tggagcagta cggcgacgag aagaaccac tgtataagta ctatgaagag
 actgggaact acctgaccaa gtatagcaaa aaggataatg gcccgtgat caagaagatc
 aagtactatg ggaacaagct gaatgccat ctggacatca cagacgatta ccctaacagt
 cgcaacaagg tgggtcaagct gtcactgaag ccatacagat tcgatgtcta tctggacaac
 ggcgtgtata aatttgtgac tgtcaagaat ctggatgtca tcaaaaagga gaactactat
 gaagtgaata gcaagtgtca cgaagaggct aaaaagctga aaaagattag caaccaggca
 gagttcatcg cctcctttta caacaacgac ctgattaaga tcaatggcga actgtatagg
 gtcacgggg tgaacaatga tctgtgaac cgcattgaag tgaatatgat tgacatcact
 taccgagagt atctggaaaa catgaatgat aagcgcccc ctcgaattat caaaacaatt
 gcctctaaga ctacagagtat caaaaagtac tcaaccgaca ttctgggaaa cctgtatgag
 gtgaagagca aaaagcacc tcagattatc aaaaagggc [SEQ ID NO: 34].

[00134] In certain embodiments, the mutant *S. aureus* Cas9 molecule comprises a N580A mutation. The nucleotide sequence encoding this mutant *S. aureus* Cas9 molecule is set forth in SEQ ID NO: 35, which is provided below:

atgaaaagga actacattct ggggctggac atcgggatta caagcgtggg gtatgggatt
 attgactatg aaacaaggga cgtgatcgac gcaggcgtca gactgttcaa ggaggccaac
 gtggaaaaca atgaggagcg gagaagcaag aggggagcca ggcgccgtgaa acgacggaga

aggcacagaa tccagaggggt gaagaaactg ctgttcgatt acaacctgct gaccgaccat
tctgagctga gtggaattaa tccttatgaa gccaggggtga aaggcctgag tcagaagctg
tcagaggaag agttttccgc agctctgctg cacctggcta agcgccgagg agtgcataac
gtcaatgagg tggaagagga caccggcaac gagctgtcta caaaggaaca gatctcacgc
aatagcaaag ctctggaaga gaagtatgtc gcagagctgc agctggaacg gctgaagaaa
gatggcgagg tgagaggggtc aattaatagg ttcaagacaa gcgactacgt caaagaagcc
aagcagctgc tgaaagtgca gaaggcttac caccagctgg atcagagctt catcgatact
tatatcgacc tgctggagac tcggagaacc tactatgagg gaccaggaga agggagcccc
ttcggatgga aagacatcaa ggaatggtac gagatgctga tgggacattg cacctatttt
ccagaagagc tgagaagcgt caagtacgct tataacgcag atctgtacaa cgccctgaat
gacctgaaca acctggtcat caccagggat gaaaacgaga aactggaata ctatgagaag
ttccagatca tcgaaaacgt gtttaagcag aagaaaaagc ctacactgaa acagattgct
aaggagatcc tgggtcaacga agaggacatc aagggtacc gggtgacaag cactggaaaa
ccagagttca ccaatctgaa agtgtatcac gatattaagg acatcacagc acggaaagaa
atcattgaga acgccgaact gctggatcag attgctaaga tcctgactat ctaccagagc
tccgaggaca tccaggaaga gctgactaac ctgaacagcg agctgacca ggaagagatc
gaacagatta gtaatctgaa ggggtacacc ggaacacaca acctgtccct gaaagctatc
aatctgattc tggatgagct gtggcataca aacgacaatc agattgcaat ctttaaccgg
ctgaagctgg tcccaaaaaa ggtggacctg agtcagcaga aagagatccc aaccacactg
gtggacgatt tcattctgtc acccgtgggtc aagcggagct tcatccagag catcaaagtg
atcaacgcca tcatcaagaa gtacggcctg ccaatgata tcattatcga gctggctagg
gagaagaaca gcaaggacgc acagaagatg atcaatgaga tgcagaaacg aaaccggcag
accaatgaac gcattgaaga gattatccga actaccggga aagagaacgc aaagtacctg
attgaaaaaa tcaagctgca cgatatgcag gagggaaagt gtctgtattc tctggaggcc
atccccctgg aggacctgct gaacaatcca ttcaactacg aggtcgatca tattatcccc
agaagcgtgt ccttcgacaa ttcctttaac aacaagggtg tgggtcaagca ggaagaggcc
tctaaaaagg gcaataggac tcctttccag tacctgtcta gttcagattc caagatctct
tacgaaacct ttaaaaagca cattctgaat ctggccaaag gaaagggccg catcagcaag
accaaaaagg agtacctgct ggaagagcgg gacatcaaca gattctccgt ccagaaggat
tttattaacc ggaatctggt ggacacaaga tacgctactc gcggcctgat gaatctgctg
cgatcctatt tccgggtgaa caatctggat gtgaaagtca agtccatcaa cggcggggtc
acatcttttc tgaggcgcaa atggaagttt aaaaaggagc gcaacaaagg gtacaagcac
catgccgaag atgctctgat tatcgcaaat gccgacttca tctttaagga gtggaaaaag
ctggacaaaag ccaagaaagt gatggagaac cagatgttcg aagagaagca ggccgaatct
atgcccgaag tcgagacaga acaggagtac aaggagattt tcactactcc tcaccagatc

aagcatatca aggatttcaa ggactacaag tactctcacc ggggtggataa aaagcccaac
 agagagctga tcaatgacac cctgtatagt acaagaaaag acgataaggg gaataccctg
 attgtgaaca atctgaacgg actgtacgac aaagataatg acaagctgaa aaagctgac
 aacaaaagtc ccgagaagct gctgatgtac caccatgatc ctacagacata tcagaaactg
 aagctgatta tggagcagta cggcgacgag aagaaccac tgtataagta ctatgaagag
 actgggaact acctgaccaa gtatagcaaa aaggataatg gccccgtgat caagaagatc
 aagtactatg ggaacaagct gaatgcccac ctggacatca cagacgatta ccctaacagt
 cgcaacaagg tgggtcaagct gtcactgaag ccatacagat tcgatgtcta tctggacaac
 ggcgtgtata aatttgtgac tgtcaagaat ctggatgtca tcaaaaagga gaactactat
 gaagtgaata gcaagtgtca cgaagaggct aaaaagctga aaaagattag caaccaggca
 gagttcatcg cctcctttta caacaacgac ctgattaaga tcaatggcga actgtatagg
 gtcacgggg tgaacaatga tctgctgaac cgcattgaag tgaatatgat tgacatcact
 taccgagagt atctggaaaa catgaatgat aagcgcccc ctcgaattat caaaacaatt
 gcctctaaga ctacagagtat caaaaagtac tcaaccgaca ttctgggaaa cctgtatgag
 gtgaagagca aaaagcacc ctagattatc aaaaagggc [SEQ ID NO: 35].

[00135] A nucleic acid encoding a Cas9 molecule can be a synthetic nucleic acid sequence. For example, the synthetic nucleic acid molecule can be chemically modified. The synthetic nucleic acid sequence can be codon optimized, e.g., at least one non-common codon or less-common codon has been replaced by a common codon. For example, the synthetic nucleic acid can direct the synthesis of an optimized messenger mRNA, e.g., optimized for expression in a mammalian expression system, e.g., described herein.

[00136] Additionally or alternatively, a nucleic acid encoding a Cas9 molecule or Cas9 polypeptide may comprise a nuclear localization sequence (NLS). Nuclear localization sequences are known in the art.

[00137] An exemplary codon optimized nucleic acid sequence encoding a Cas9 molecule of *S. pyogenes* is set forth in SEQ ID NO: 26, which is provided below:

atggataaaa agtacagcat cgggctggac atcgggtacaa actcagtggg gtggggccgtg
 attacggacg agtacaaggt accctccaaa aaattttaaag tgctgggtaa cacggacaga
 cactctataa agaaaaatct tattggagcc ttgctgttcg actcaggcga gacagccgaa
 gccacaaggt tgaagcggac cgccaggagg cggatatcca ggagaaagaa ccgcatatgc
 tacctgcaag aaatcttcag taacgagatg gcaaagggtg acgatagctt tttccatcgc
 ctggaagaat cctttcttgt tgaggaagac aagaagcacg aacggcacc cactcttggc
 aatattgtcg acgaagtggc atatcacgaa aagtaccgga ctatctacca cctcaggaag

aagctggtgg actctaccga taaggcggac ctgagactta tttatattggc actcgcccac
atgattaaat ttagaggaca tttcttgatc gagggcgacc tgaacccgga caacagtgac
gtcgataagc tgttcatcca acttggtgcag acctacaatc aactgttcga agaaaaccct
ataaatgctt caggagtcga cgctaaagca atcctgtccg cgcgcctctc aaaatctaga
agacttgaga atctgattgc tcagttgccc ggggaaaaga aaaatggatt gtttggaac
ctgatcgccc tcagtctcgg actgaccca aatttcaaaa gtaacttcga cctggccgaa
gacgctaagc tccagctgtc caaggacaca tacgatgacg acctcgacaa tctgctggcc
cagattgggg atcagtagc cgatctcttt ttggcagcaa agaacctgtc cgacgccatc
ctgttgagcg atatcttgag agtgaacacc gaaattacta aagcaccct tagcgcactc
atgatcaagc ggtacgacga gcatcatcag gatctgacct tgctgaaggc tcttgtaggg
caacagctcc ccgaaaaata caaggaaatc ttctttgacc agagcaaaaa cggctacgct
ggctatatag atggtggggc cagtcaggag gaattctata aattcatcaa gccattctc
gagaaaaatgg acggcacaga ggagttgctg gtcaaaacta acagggagga cctgctgcgg
aagcagcgga cctttgacaa cgggtctatc cccaccaga ttcactctggg cgaactgcac
gcaatcctga ggaggcagga ggatttttat ctttttctta aagataaccg cgagaaaaata
gaaaagattc ttacattcag gatcccgtac tacgtgggac ctctcgcccg gggcaattca
cggtttgctt ggatgacaag gaagtcagag gagactatta caccttgga cttcgaagaa
gtggtggaca aggggtgcac tgcccagtct ttcactcgagc ggatgacaaa ttttgacaag
aacctcccta atgagaaggt gctgcccata cattctctgc tctacgagta ctttaccgtc
tacaatgaac tgactaaagt caagtacgtc accgaggga tgagggaagc ggcattcctt
agtggagaac agaagaaggc gattgtagac ctggtgttca agaccaacag gaaggtagt
gtgaagcaac ttaaagaaga ctactttaag aagatcgaat gttttgacag tgtggaaatt
tcaggggttg aagaccgctt caatgcgtca ttggggactt accatgatct tctcaagatc
ataaaggaca aagacttcct ggacaacgaa gaaaatgagg atattctcga agacatcgtc
ctcaccctga cctgttcga agacaggga atgatagaag agcgcttgaa aacctatgcc
cacctcttcg acgataaagt tatgaagcag ctgaagcgca ggagatacac aggatgggga
agattgtcaa ggaagctgat caatggaatt agggataaac agagtggcaa gaccatactg
gatttcctca aatctgatgg cttcgccaat aggaacttca tgcaactgat tcacgatgac
tctcttacct tcaaggagga cattcaaaag gctcagggtga gcgggcaggg agactccctt
catgaacaca tcgcgaattt ggcagggttc cccgctatta aaaaggcat ccttcaaact
gtcaagggtg tggatgaatt ggtcaaggta atgggcagac ataagccaga aaatattgtg
atcgagatgg cccgcgaaaa ccagaccaca cagaagggcc agaaaaatag tagagagcgg
atgaagagga tcgaggagg catcaaagag ctgggatctc agattctcaa agaacacccc
gtagaaaaca cacagctgca gaacgaaaaa ttgtacttgt actatctgca gaacggcaga
gacatgtacg tcgaccaaga acttgatatt aatagactgt ccgactatga cgtagacat

atcgtgcccc agtccttcct gaaggacgac tccattgata acaaagtctt gacaagaagc
 gacaagaaca ggggtaaaag tgataatgtg cctagcgagg aggtggtgaa aaaaatgaag
 aactactggc gacagctgct taatgcaaag ctcatcac aacggaagtt cgataatctg
 acgaaagcag agagaggtgg cttgtctgag ttggacaagg cagggtttat taagcggcag
 ctggtggaac ctaggcagat cacaagcac gtggcgcaga ttttgacag ccgcatgaac
 aaaaaatacg acgaaaatga taaactgata cgagaggtca aagttatcac gctgaaaagc
 aagctggtgt ccgattttcg gaaagacttc cagttctaca aagttcgcga gattaataac
 taccatcatg ctacagatgc gtacctgaac gctgttgcg ggaccgcctt gataaagaag
 tacccaaagc tggaatccga gttcgtatac ggggattaca aagtgtacga tgtgaggaaa
 atgatagcca agtccgagca ggagattgga aaggccacag ctaagtactt cttttattct
 aacatcatga atttttttaa gacggaaatt accctggcca acggagagat cagaaagcgg
 ccccttatag agacaaatgg tgaaacaggt gaaatcgtct gggataaggg cagggtttc
 gctactgtga ggaaggtgct gagtatgcca caggtaaata tcgtgaaaaa aaccgaagta
 cagaccggag gattttccaa ggaaagcatt ttgcctaaaa gaaactcaga caagctcatc
 gcccgcaaga aagattggga ccctaagaaa tacgggggat ttgactcacc caccgtagcc
 tattctgtgc tgggtgtagc taaggtggaa aaaggaaagt ctaagaagct gaagtccgtg
 aaggaaactt tgggaatcac tatcatggaa agatcatcct ttgaaaagaa ccctatcgat
 ttcttgagg ctaagggtta caaggaggtc aagaaagacc tcatcattaa actgcaaaaa
 tactctctct tcgagctgga aaatggcagg aagagaatgt tggccagcgc cggagagctg
 caaaaggga acgagcttgc tctgccctcc aaatatgtta attttctcta tctcgcttcc
 cactatgaaa agctgaaagg gtctcccgaa gataacgagc agaagcagct gttcgtcgaa
 cagcacaagc actatctgga tgaaataatc gaacaaataa gcgagttcag caaaagggtt
 atcctggcgg atgctaattt ggacaaagta ctgtctgctt ataacaagca ccgggataag
 cctattaggg aacaagccga gaataaatt cacctcttta cactcacgaa tctcggagcc
 cccgccgct tcaaatactt tgatacgact atcgaccgga aacggtatac cagtaccaa
 gaggtcctcg atgccaccct catccaccag tcaattactg gcctgtacga

aacacggatcgacctctctc aactgggcgg cgactag [SEQ ID NO: 26].

[00138] The corresponding amino acid sequence of an *S. pyogenes* Cas9 molecule is set forth in SEQ ID NO: 27, which is provided below:

MDKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLK
 RTARRRYTRRKNRICYLQEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHE
 KYPTIYHLRKKLVSTDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVKLFIQLVQTYNQL
 FEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAE
 DAKLQLSKDITYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY

DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELL
 VKLNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFY PFLKDNREKIEKILTFRIPYYVGPL
 ARGNSRFAWMTRKSEETITPWNFEEVVDKGASQSFIERMTNFDKNLPNEKVLPHSLLYEYFTV
 YNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVED
 RFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTTLTFEDREMIEERLKTYAHLFDDKVMKQ
 LKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVSG
 QGDSLHEHIANLAGSPAIIKKGILQTVKVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRER
 MKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQS
 FLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLITQRKFDNLTKAERGGLSE
 LDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFYKV
 REINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEIGKATAKYFFYS
 NIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIIVKKTEVQTGGF
 SKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKSKKLKSVKELLGITIME
 RSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLEFELENGRKRMLASAGELQKGNELALPSKYVNF
 LYLASHYEKLKGGSPEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVLADANLDKVL SAYNKH RDK
 PIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQL
 GGD [SEQ ID NO: 27]

[00139] Exemplary codon optimized nucleic acid sequences encoding a Cas9 molecule of *S. aureus*, and optionally containing nuclear localization sequences (NLSs), are set forth in SEQ ID NOs: 28-32, 43, and 44, which are provided below. Another exemplary codon optimized nucleic acid sequence encoding a Cas9 molecule of *S. aureus* comprises the nucleotides 1293-4451 of SEQ ID NO: 83.

[00140] SEQ ID NO: 28 is set forth below:

atgaaaagga actacattct ggggctggac atcgggatta caagcgtggg gtatgggatt
 attgactatg aaacaaggga cgtgatcgac gcaggcgtca gactgttcaa ggaggccaac
 gtggaaaaca atgaggggacg gagaagcaag aggggagcca ggcgcctgaa acgacggaga
 aggcacagaa tccagagggt gaagaaactg ctgttcgatt acaacctgct gaccgaccat
 tctgagctga gtggaattaa tccttatgaa gccagggtga aaggcctgag tcagaagctg
 tcagaggaag agttttccgc agctctgctg cacctggcta agcgccgagg agtgcataac
 gtcaatgagg tggaagagga caccggcaac gagctgtcta caaaggaaca gatctcacgc
 aatagcaaag ctctggaaga gaagtatgtc gcagagctgc agctggaacg gctgaagaaa
 gatggcgagg tgagagggtc aattaatagg ttcaagacaa gcgactacgt caaagaagcc
 aagcagctgc tgaaagtgca gaaggcttac caccagctgg atcagagctt catcgatact
 tatatcgacc tgctggagac tcggagaacc tactatgagg gaccaggaga agggagcccc

ttcggatgga aagacatcaa ggaatggtac gagatgctga tgggacattg cacctatttt
ccagaagagc tgagaagcgt caagtacgct tataacgcag atctgtacaa cgccctgaat
gacctgaaca acctggtcat caccagggat gaaaacgaga aactggaata ctatgagaag
ttccagatca tcgaaaacgt gtttaagcag aagaaaaagc ctacactgaa acagattgct
aaggagatcc tgggtcaacga agaggacatc aagggtacc ggggtgacaag cactggaaaa
ccagagttca ccaatctgaa agtgtatcac gatattaagg acatcacagc acggaaagaa
atcattgaga acgccgaact gctggatcag attgctaaga tcctgactat ctaccagagc
tccgaggaca tccaggaaga gctgactaac ctgaacagcg agctgaccca ggaagagatc
gaacagatta gtaatctgaa ggggtacacc ggaacacaca acctgtccct gaaagctatc
aatctgattc tggatgagct gtggcataca aacgacaatc agattgcaat ctttaaccgg
ctgaagctgg tccccaaaaa ggtggacctg agtcagcaga aagagatccc aaccacactg
gtggacgatt tcattctgtc acccgtggtc aagcggagct tcatccagag catcaaagtg
atcaacgcca tcatcaagaa gtacggcctg ccaatgata tcattatcga gctggctagg
gagaagaaca gcaaggacgc acagaagatg atcaatgaga tgcagaaacg aaaccggcag
accaatgaac gcattgaaga gattatccga actaccgga aagagaacgc aaagtacctg
attgaaaaaa tcaagctgca cgatatgcag gagggaaagt gtctgtattc tctggaggcc
tccccctgg aggacctgct gaacaatcca ttcaactacg aggtcgatca tattatcccc
agaagcgtgt ccttcgacaa ttcctttaac aacaagggtc tgggtcaagca ggaagagaac
tctaaaaagg gcaataggac tcctttccag tacctgtcta gttcagattc caagatctct
tacgaaacct ttaaaaagca cattctgaat ctggccaaag gaaagggccg catcagcaag
accaaaaagg agtacctgct ggaagagcgg gacatcaaca gattctccgt ccagaaggat
tttattaacc ggaatctggt ggacacaaga tacgctactc gcggcctgat gaatctgctg
cgatcctatt tccgggtgaa caatctggat gtgaaagtca agtccatcaa cggcgggttc
acatcttttc tgaggcgcaa atggaagttt aaaaaggagc gcaacaaagg gtacaagcac
catgccgaag atgctctgat tatcgcaaat gccgacttca tctttaagga gtggaaaaag
ctggacaaaag ccaagaaagt gatggagaac cagatgttcg aagagaagca ggccgaatct
atgcccgaag tcgagacaga acaggagtac aaggagattt tcatcactcc tcaccagatc
aagcatatca aggatttcaa ggactacaag tactctcacc ggggtgataa aaagcccaac
agagagctga tcaatgacac cctgtatagt acaagaaaag acgataaggg gaataacctg
attgtgaaca atctgaacgg actgtacgac aaagataatg acaagctgaa aaagctgatc
aacaaaagtc ccgagaagct gctgatgtac caccatgatc ctacagacata tcagaaactg
aagctgatta tggagcagta cggcgacgag aagaaccac tgtataagta ctatgaagag
actgggaact acctgaccaa gtatagcaaa aaggataatg gccccgtgat caagaagatc
aagtactatg ggaacaagct gaatgcccat ctggacatca cagacgatta ccctaacagt
cgcaacaagg tgggtcaagct gtcactgaag ccatacagat tcgatgtcta tctggacaac

ggcgtgtata aatttgtgac tgtcaagaat ctggatgtca tcaaaaagga gaactactat
 gaagtgaata gcaagtgtga cgaagaggct aaaaagctga aaaagattag caaccaggca
 gagttcatcg cctcctttta caacaacgac ctgattaaga tcaatggcga actgtatagg
 gtcacgggg tgaacaatga tctgctgaac cgcattgaag tgaatatgat tgacatcact
 taccgagagt atctggaaaa catgaatgat aagcgcccc ctcgaattat caaaacaatt
 gcctctaaga ctacagagtat caaaaagtac tcaaccgaca ttctgggaaa cctgtatgag
 gtgaagagca aaaagcacco tcagattatc aaaaagggc [SEQ ID NO: 28]

[00141] SEQ ID NO: 29 is set forth below.

atgaagcgga actacatcct gggcctggac atcggcacatca ccagcgtggg ctacggcatc
 atcgactacg agacacggga cgtgatcgat gccggcgtgc ggctgttcaa agaggccaac
 gtggaaaaca acgagggcag gcggagcaag agaggcgcca gaaggctgaa gcggcgagg
 cggcatagaa tccagagagt gaagaagctg ctgttcgact acaacctgct gaccgaccac
 agcgagctga gcggcatcaa cccctacgag gccagagtga agggcctgag ccagaagctg
 agcgaggaag agttctctgc cgccctgctg cacctggcca agagaagagg cgtgcacaac
 gtgaacgagg tggaagagga caccggcaac gagctgtcca ccaaagagca gatcagccgg
 aacagcaagg ccctggaaga gaaatacgtg gccgaactgc agctggaacg gctgaagaaa
 gacggcgaag tgcggggcag catcaacaga ttcaagacca gcgactacgt gaaagaagcc
 aaacagctgc tgaaggtgca gaaggcctac caccagctgg accagagctt catcgacacc
 tacatcgacc tgctggaac ccggcggacc tactatgagg gacctggcga gggcagcccc
 ttgggctgga aggacatcaa agaattggtac gagatgctga tggggccactg cacctacttc
 cccgaggaac tgcggagcgt gaagtacgcc tacaacgccg acctgtacaa cgccctgaac
 gacctgaaca atctcgtgat caccaggac gagaacgaga agctggaata ttacgagaag
 ttccagatca tcgagaacgt gttcaagcag aagaagaagc ccacctgaa gcagatcgcc
 aaagaaatcc tcgtgaacga agaggatatt aagggtaca gagtgaccag caccggcaag
 cccgagttca ccaacctgaa ggtgtaccac gacatcaagg acattaccgc ccggaaagag
 attattgaga acgccgagct gctggatcag attgccaaaga tcctgaccat ctaccagagc
 agcgaggaca tccaggaaga actgaccaat ctgaactccg agctgacca ggaagagatc
 gagcagatct ctaatctgaa gggctatacc ggcaccaca acctgagcct gaaggccatc
 aacctgatcc tggacgagct gtggcacacc aacgacaacc agatcgctat cttcaaccgg
 ctgaagctgg tgcccaagaa ggtggacctg tcccagcaga aagagatccc caccacctg
 gtggacgact tcatcctgag cccgctcgtg aagagaagct tcatccagag catcaaagtg
 atcaacgcca tcatcaagaa gtacggcctg cccaacgaca tcattatcga gctggcccg
 gagaagaact ccaaggacgc ccagaaaatg atcaacgaga tgcagaagcg gaaccggcag
 accaacgagc ggatcgagga aatcatccgg accaccggca aagagaacgc caagtacctg

atcgagaaga tcaagctgca cgacatgcag gaaggcaagt gcctgtacag cctggaagcc
 atccctctgg aagatctgct gaacaacccc ttcaactatg aggtggacca catcatcccc
 agaagcgtgt ccttcgacaa cagcttcaac aacaagggtgc tcgtgaagca ggaagaaaac
 agcaagaagg gcaaccggac ccattccag tacctgagca gcagcgacag caagatcagc
 tacgaaacct tcaagaagca catcctgaat ctggccaagg gcaagggcag aatcagcaag
 accaagaaaag agtatctgct ggaagaacgg gacatcaaca ggttctccgt gcagaaagac
 ttcatacaacc ggaacctggt ggataccaga tacgccacca gaggcctgat gaacctgctg
 cggagctact tcagagtga caacctggac gtgaaagtga agtccatcaa tggcggcttc
 accagctttc tgcggcggaa gtggaagttt aagaaagagc ggaacaaggg gtacaagcac
 cagccgagg acgccctgat cattgccaac gccgatttca tcttcaaaga gtggaagaaa
 ctggacaagg ccaaaaaagt gatggaaaac cagatgttcg aggaaaagca ggccgagagc
 atgcccgaga tcgaaaccga gcaggagtac aaagagatct tcatcacccc ccaccagatc
 aagcacatta aggacttcaa ggactacaag tacagccacc ggggtggacaa gaagcctaata
 agagagctga ttaacgacac cctgtactcc acccggaagg acgacaaggg caacaccctg
 atcgtgaaca atctgaacgg cctgtacgac aaggacaatg acaagctgaa aaagctgatc
 aacaagagcc ccgaaaagct gctgatgtac caccacgacc ccagaccta ccagaaactg
 aagctgatta tggaacagta cggcgacgag aagaatcccc tgtacaagta ctacgaggaa
 accgggaact acctgaccaa gtactccaaa aaggacaacg gccccgtgat caagaagatt
 aagtattacg gcaacaaact gaacgcccat ctggacatca ccgacgacta ccccaacagc
 agaaacaagg tcgtgaagct gtccctgaag cctacagat tcgacgtgta cctggacaat
 ggcgtgtaca agttcgtgac cgtgaagaat ctggatgtga tcaaaaaaga aaactactac
 gaagtgaata gcaagtgtca tgaggaagct aagaagctga agaagatcag caaccaggcc
 gagtttatcg cctccttcta caacaacgat ctgatcaaga tcaacggcga gctgtataga
 gtgatcggcg tgaacaacga cctgctgaac cggatcgaag tgaacatgat cgacatcacc
 taccgagagt acctggaaaa catgaacgac aagaggcccc ccaggatcat taagacaatc
 gcctccaaga ccagagcat taagaagtac agcacagaca ttctgggcaa cctgtatgaa
 gtgaaatcta agaagcacc tcagatcatc aaaaagggc [SEQ ID NO: 29]

[00142] SEQ ID NO: 30 is set forth below.

atgaagcgca actacatcct cggactggac atcggcatta cctccgtggg atacggcatc
 atcgattacg aaactaggga tgtgatcgac gctggagtca ggctgttcaa agaggcgaac
 gtggagaaca acgaggggcg gcgctcaaag aggggggccc gccggctgaa gcgcccgcgc
 agacatagaa tccagcgct gaagaagctg ctgttcgact acaaccttct gaccgaccac
 tccgaacttt ccggcatcaa cccatatgag gctagagtga agggattgtc ccaaaagctg
 tccgaggaag agttctccgc cgcgttgctc cacctcgcca agcgcagggg agtgcacaat

gtgaacgaag tggaagaaga taccggaaac gagctgtcca ccaaggagca gatcagccgg
aactccaagg ccctggaaga gaaatacgtg gcggaactgc aactggagcg gctgaagaaa
gacggagaag tgcgcggctc gatcaaccgc ttcaagacct cggactacgt gaaggaggcc
aagcagctcc tgaaagtgc aaaggcctat caccaacttg accagtcctt tatcgatacc
tacatcgatc tgctcgagac tcggcggact tactacgagg gtccagggga gggctcccca
tttggttgga aggatattaa ggagtggtag gaaatgctga tgggacactg cacatacttc
cctgaggagc tgcggagcgt gaaatacgc tacaacgcag acctgtacaa cgcgctgaac
gacctgaaca atctcgtgat caccggggac gagaacgaaa agctcgagta ttacgaaaag
ttccagatta ttgagaacgt gttcaaacag aagaagaagc cgacactgaa gcagattgcc
aaggaaatcc tcgtgaacga agaggacatc aagggtatc gagtgacctc aacgggaaag
ccggagttca ccaatctgaa ggtctaccac gacatcaaag acattaccgc ccggaaggag
atcattgaga acgcggagct gttggaccag attgcgaaga ttctgaccat ctaccaatcc
tccgaggata ttcaggaaga actcaccaac ctcaacagcg aactgacca ggaggagata
gagcaaatct ccaacctgaa gggctacacc ggaactcata acctgagcct gaaggccatc
aacttgatcc tggacgagct gtggcacacc aacgataacc agatcgctat tttcaatcgg
ctgaagctgg tccccaagaa agtggacctc tcacaacaaa aggagatccc tactaccctt
gtggacgatt tcattctgtc ccccggtggtc aagagaagct tcatacagtc aatcaaagtg
atcaatgcca ttatcaagaa atacggtctg cccaacgaca ttatcattga gctcgccgcg
gagaagaact cgaaggacgc ccagaagatg attaacgaaa tgcagaagag gaaccgacag
actaacgaac ggatcgaaga aatcatccgg accaccggga aggaaaacgc gaagtacctg
atcgaaaaga tcaagctcca tgacatgcag gaaggaaagt gtctgtactc gctggaggcc
attccgctgg aggacttgct gaacaaccct ttaactacg aagtggatca tatcattccg
aggagcgtgt cattcgacaa ttccttcaac aacaagggtc tcgtgaagca ggaggaaaac
tcgaagaagg gaaaccgcac gccgttccag tacctgagca gcagcgactc caagatttcc
tacgaaacct tcaagaagca catcctcaac ctggcaaagg ggaagggtcg catctccaag
accaagaagg aatatctgct ggaagaaaga gacatcaaca gattctccgt gcaaaaggac
ttcatcaacc gcaacctcgt ggatactaga tacgctactc ggggtctgat gaacctcctg
agaagctact ttagagtga caatctggac gtgaagggtca agtcgattaa cggaggtttc
acctccttcc tgcggcgcaa gtggaagttc aagaaggaaac ggaacaaggg ctacaagcac
cacgccgagg acgccctgat cattgccaac gccgacttca tcttcaaaga atggaagaaa
cttgacaagg ctaagaagggt catggaaaac cagatgttcg aagaaaagca ggccgagtct
atgcctgaaa tcgagactga acaggagtac aaggaaatct ttattacgcc acaccagatc
aaacacatca aggatttcaa ggattacaag tactcacatc gcgtggacaa aaagccgaac
agggaactga tcaacgacac cctctactcc acccggaagg atgacaaagg gaataccctc
atcgtcaaca accttaacgg cctgtacgac aaggacaacg ataagctgaa gaagctcatt

aacaagtgcg ccgaaaagtt gctgatgtac caccacgacc ctgagactta ccagaagctc
aagctgatca tggagcagta tggggacgag aaaaacccgt tgtacaagta ctacgaagaa
actgggaatt atctgactaa gtactccaag aaagataacg gccccgtgat taagaagatt
aagtactacg gcaacaagct gaacgccccat ctggacatca ccgatgacta ccctaattcc
cgcaacaagg tcgtcaagct gagcctcaag ccctaccggt ttgatgtgta ccttgacaat
ggagtgtaca agttcgtgac tgtgaagaac cttgacgtga tcaagaagga gaactactac
gaagtcaact ccaagtgtga cgaggaagca aagaagttga agaagatctc gaaccaggcc
gagttcattg cctccttcta taacaacgac ctgattaaga tcaacggcga actgtaccgc
gtcattggcg tgaacaacga tctcctgaac cgcacgaag tgaacatgat cgacatcact
taccgggaat acctggagaa tatgaacgac aagcgcccg cccggatcat taagactatc
gcctcaaaga ccagtcgat caagaagtac agcaccgaca tcctgggcaa cctgtacgag
gtcaaatcga agaagcacc ccagatcatc aagaaggga [SEQ ID NO: 30]

[00143] SEQ ID NO: 31 is set forth below.

ATGGCCCCAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCAAGCGGAACTACAT
CCTGGGCCTGGACATCGGCATCACCAGCGTGGGCTACGGCATCATCGACTACGAGACACGGGACG
TGATCGATGCCGGCGTGCGGCTGTTCAAAGAGGCCAACGTGGAAAACAACGAGGGCAGGCGGAGC
AAGAGAGGCGCCAGAAGGCTGAAGCGGCGGAGGCGGCATAGAATCCAGAGAGTGAAGAAGCTGCT
GTTCTGACTACAACCTGCTGACCGACCACAGCGAGCTGAGCGGCATCAACCCCTACGAGGCCAGAG
TGAAGGGCCTGAGCCAGAAGCTGAGCGAGGAAGAGTTCTCTGCCGCCCTGCTGCACCTGGCCAAG
AGAAGAGGCGTGACAACGTGAACGAGGTGGAAGAGGACACCGGCAACGAGCTGTCCACCAGAGA
GCAGATCAGCCGGAACAGCAAGGCCCTGGAAGAGAAATACGTGGCCGAACTGCAGCTGGAACGGC
TGAAGAAAGACGGCGAAGTGCGGGGCAGCATCAACAGATTCAAGACCAGCGACTACGTGAAAGAA
GCCAAACAGCTGCTGAAGGTGCAGAAGGCCCTACCACCAGCTGGACCAGAGCTTCATCGACACCTA
CATCGACCTGCTGGAAACCCGGCGGACCTACTATGAGGGACCTGGCGAGGGCAGCCCCTTCGGCT
GGAAGGACATCAAAGAATGGTACGAGATGCTGATGGGCCACTGCACCTACTTCCCCGAGGAACTG
CGGAGCGTGAAGTACGCCTACAACGCCGACCTGTACAACGCCCTGAACGACCTGAACAATCTCGT
GATCACCAGGGACGAGAACGAGAAGCTGGAATATTACGAGAAGTTCCAGATCATCGAGAACGTGT
TCAAGCAGAAGAAGAAGCCACCCTGAAGCAGATCGCCAAAGAAATCCTCGTGAACGAAGAGGAT
ATTAAGGGCTACAGAGTGACCAGCACCGGCAAGCCCGAGTTCACCAACCTGAAGGTGTACCACGA
CATCAAGGACATTACCGCCCGGAAAGAGATTATTGAGAACGCCGAGCTGCTGGATCAGATTGCCA
AGATCCTGACCATCTACCAGAGCAGCGAGGACATCCAGGAAGAACTGACCAATCTGAACTCCGAG
CTGACCCAGGAAGAGATCGAGCAGATCTCTAATCTGAAGGGCTATACCGGCACCCACAACCTGAG
CCTGAAGGCCATCAACCTGATCCTGGACGAGCTGTGGCACACCAACGACAACCAGATCGCTATCT
TCAACCGGCTGAAGCTGGTGCCCAAGAAGGTGGACCTGTCCCAGCAGAAAGAGATCCCCACCACC

CTGGTGGACGACTTCATCCTGAGCCCCGTCGTGAAGAGAAGCTTCATCCAGAGCATCAAAGTGAT
CAACGCCATCATCAAGAAGTACGGCCTGCCCAACGACATCATTATCGAGCTGGCCCGCGAGAAGA
ACTCCAAGGACGCCCAGAAAATGATCAACGAGATGCAGAAGCGGAACCGGCAGACCAACGAGCGG
ATCGAGGAAATCATCCGGACCACCGGCAAAGAGAACGCCAAGTACCTGATCGAGAAGATCAAGCT
GCACGACATGCAGGAAGGCAAGTGCCTGTACAGCCTGGAAGCCATCCCTCTGGAAGATCTGCTGA
ACAACCCCTTCAACTATGAGGTGGACCACATCATCCCCAGAAGCGTGTCTTCGACAACAGCTTC
AACACAAGGTGCTCGTGAAGCAGGAAGAAAACAGCAAGAAGGGCAACCGGACCCCATTCAGTA
CCTGAGCAGCAGCGACAGCAAGATCAGCTACGAAACCTTCAAGAAGCACATCCTGAATCTGGCCA
AGGGCAAGGGCAGAATCAGCAAGACCAAGAAAGAGTATCTGCTGGAAGAACGGGACATCAACAGG
TTCTCCGTGCAGAAAGACTTCATCAACCGGAACCTGGTGGATACCAGATACGCCACCAGAGGCCT
GATGAACCTGCTGCGGAGCTACTTCAGAGTGAACAACCTGGACGTGAAAGTGAAGTCCATCAATG
GCGGCTTCACCAGCTTTCTGCGGCGGAAGTGGAGTTTAAAGAAAGAGCGGAACAAGGGGTACAAG
CACCACGCCGAGGACGCCCTGATCATTGCCAACGCCGATTTTCATCTTCAAAGAGTGAAGAAACT
GGACAAGGCCAAAAAAGTGATGGAAAACCAGATGTTTCGAGGAAAGGCAGGCCGAGAGCATGCCCCG
AGATCGAAACCGAGCAGGAGTACAAAGAGATCTTCATCACCCCCCACCAGATCAAGCACATTAAG
GACTTCAAGGACTACAAGTACAGCCACCGGGTGGACAAGAAGCCTAATAGAGAGCTGATTAACGA
CACCCTGTACTCCACCCGGAAGGACGACAAGGGCAACACCCTGATCGTGAACAATCTGAACGGCC
TGTACGACAAGGACAATGACAAGCTGAAAAAGCTGATCAACAAGAGCCCCGAAAAGCTGCTGATG
TACCACCACGACCCCCAGACCTACCAGAACTGAAGCTGATTATGGAACAGTACGGCGACGAGAA
GAATCCCCTGTACAAGTACTACGAGGAAACCGGGAACCTACCTGACCAAGTACTCCAAAAGGACA
ACGGCCCCGTGATCAAGAAGATTAAGTATTACGGCAACAACTGAACGCCCATCTGGACATCACC
GACGACTACCCCAACAGCAGAAACAAGGTCTGAAGCTGTCCCTGAAGCCCTACAGATTCGACGT
GTACCTGGACAATGGCGTGTACAAGTTCGTGACCGTGAAGAATCTGGATGTGATCAAAAAGAAA
ACTACTACGAAGTGAATAGCAAGTGCTATGAGGAAGCTAAGAAGCTGAAGAAGATCAGCAACCAG
GCCGAGTTTATCGCCTCCTTCTACAACAACGATCTGATCAAGATCAACGGCGAGCTGTATAGAGT
GATCGGCGTGAACAACGACCTGCTGAACCGGATCGAAGTGAACATGATCGACATCACCTACCGCG
AGTACCTGGAAAACATGAACGACAAGAGGCCCCCAGGATCATTAAGACAATCGCCTCCAAGACC
CAGAGCATTAAGAAGTACAGCACAGACATTCTGGGCAACCTGTATGAAGTGAATCTAAGAAGCA
CCCTCAGATCATCAAAAAGGGCAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGA
AAAAG [SEQ ID NO: 31]

[00144] SEQ ID NO: 32 is set forth below.

ACCGGTGCCA CCATGTACCC ATACGATGTT CCAGATTACG CTTCGCCGAA GAAAAAGCGC
AAGGTCGAAG CGTCCATGAA AAGGAACTAC ATTCTGGGGC TGGACATCGG GATTACAAGC
GTGGGGTATG GGATTATTGA CTATGAAACA AGGGACGTGA TCGACGCAGG CGTCAGACTG

TTCAAGGAGG CCAACGTGGA AAACAATGAG GGACGGAGAA GCAAGAGGGG AGCCAGGCGC
 CTGAAACGAC GGAGAAGGCA CAGAATCCAG AGGGTGAAGA AACTGCTGTT CGATTACAAC
 CTGCTGACCG ACCATTCTGA GCTGAGTGGA ATTAATCCTT ATGAAGCCAG GGTGAAAGGC
 CTGAGTCAGA AGCTGTCAGA GGAAGAGTTT TCCGCAGCTC TGCTGCACCT GGCTAAGCGC
 CGAGGAGTGC ATAACGTCAA TGAGGTGGAA GAGGACACCG GCAACGAGCT GTCTACAAAG
 GAACAGATCT CACGCAATAG CAAAGCTCTG GAAGAGAAGT ATGTGCGAGA GCTGCAGCTG
 GAACGGCTGA AGAAAGATGG CGAGGTGAGA GGGTCAATTA ATAGGTTCAA GACAAGCGAC
 TACGTCAAAG AAGCCAAGCA GCTGCTGAAA GTGCAGAAGG CTTACCACCA GCTGGATCAG
 AGCTTCATCG ATACTTATAT CGACCTGCTG GAGACTCGGA GAACCTACTA TGAGGGACCA
 GGAGAAGGGA GCCCCTTCGG ATGGAAAGAC ATCAAGGAAT GGTACGAGAT GCTGATGGGA
 CATTGCACCT ATTTTCCAGA AGAGCTGAGA AGCGTCAAGT ACGCTTATAA CGCAGATCT
 TACAACGCCC TGAATGACCT GAACAACCTG GTCATCACCA GGGATGAAAA CGAGAAACTG
 GAATACTATG AGAAGTTCCA GATCATCGAA AACGTGTTTA AGCAGAAGAA AAAGCCTACA
 CTGAAACAGA TTGCTAAGGA GATCCTGGTC AACGAAGAGG ACATCAAGGG CTACCGGGTG
 ACAAGCACTG GAAAACCAGA GTTCACCAAT CTGAAAGTGT ATCACGATAT TAAGGACATC
 ACAGCACGGA AAGAAATCAT TGAGAACGCC GAACTGCTGG ATCAGATTGC TAAGATCCTG
 ACTATCTACC AGAGCTCCGA GGACATCCAG GAAGAGCTGA CTAACCTGAA CAGCGAGCTG
 ACCCAGGAAG AGATCGAACA GATTAGTAAT CTGAAGGGGT ACACCGGAAC ACACAACCTG
 TCCCTGAAAG CTATCAATCT GATTCTGGAT GAGCTGTGGC ATACAAACGA CAATCAGATT
 GCAATCTTTA ACCGGCTGAA GCTGGTCCCA AAAAAGGTGG ACCTGAGTCA GCAGAAAGAG
 ATCCCAACCA CACTGGTGGA CGATTTCATT CTGTCACCCG TGGTCAAGCG GAGCTTCATC
 CAGAGCATCA AAGTGATCAA CGCCATCATC AAGAAGTACG GCCTGCCCCA TGATATCATT
 ATCGAGCTGG CTAGGGAGAA GAACAGCAAG GACGCACAGA AGATGATCAA TGAGATGCAG
 AAACGAAACC GGCAGACCAA TGAACGCATT GAAGAGATTA TCCGAACTAC CGGGAAAGAG
 AACGCAAAGT ACCTGATTGA AAAAATCAAG CTGCACGATA TGCAGGAGGG AAAGTGTCTG
 TATTCTCTGG AGGCCATCCC CCTGGAGGAC CTGCTGAACA ATCCATTCAA CTACGAGGTC
 GATCATATTA TCCCCAGAAG CGTGTCTTTC GACAATTCCT TTAACAACAA GGTGCTGGTC
 AAGCAGGAAG AGAACTCTAA AAAGGGCAAT AGGACTCCTT TCCAGTACCT GTCTAGTTCA
 GATTCCAAGA TCTCTTACGA AACCTTTAAA AAGCACATTC TGAATCTGGC CAAAGGAAAG
 GGCCGCATCA GCAAGACCAA AAAGGAGTAC CTGCTGGAAG AGCGGGACAT CAACAGATTCT
 TCCGTCCAGA AGGATTTTAT TAACCGGAAT CTGGTGGACA CAAGATACGC TACTCGCGGC
 CTGATGAATC TGCTGCGATC CTATTTCCGG GTGAACAATC TGGATGTGAA AGTCAAGTCC
 ATCAACGGCG GGTTACATC TTTTCTGAGG CGCAAATGGA AGTTTAAAAA GGAGCGCAAC
 AAAGGGTACA AGCACCATGC CGAAGATGCT CTGATTATCG CAAATGCCGA CTTTATCTTT
 AAGGAGTGGA AAAAGCTGGA CAAAGCCAAG AAAGTGATGG AGAACCAGAT GTTCGAAGAG

AAGCAGGCCG AATCTATGCC CGAAATCGAG ACAGAACAGG AGTACAAGGA GATTTTCATC
 ACTCCTCACC AGATCAAGCA TATCAAGGAT TTCAAGGACT ACAAGTACTC TCACCGGGTG
 GATAAAAAGC CCAACAGAGA GCTGATCAAT GACACCCTGT ATAGTACAAG AAAAGACGAT
 AAGGGGAATA CCCTGATTGT GAACAATCTG AACGGACTGT ACGACAAAGA TAATGACAAG
 CTGAAAAAGC TGATCAACAA AAGTCCCGAG AAGCTGCTGA TGTACCACCA TGATCCTCAG
 ACATATCAGA AACTGAAGCT GATTATGGAG CAGTACGGCG ACGAGAAGAA CCCACTGTAT
 AAGTACTATG AAGAGACTGG GAACTACCTG ACCAAGTATA GCAAAAAGGA TAATGGCCCC
 GTGATCAAGA AGATCAAGTA CTATGGGAAC AAGCTGAATG CCCATCTGGA CATCACAGAC
 GATTACCCTA ACAGTCGCAA CAAGGTGGTC AAGCTGTAC TGAAGCCATA CAGATTCGAT
 GTCTATCTGG ACAACGGCGT GTATAAATTT GTGACTGTCA AGAATCTGGA TGTCAATCAA
 AAGGAGAACT ACTATGAAGT GAATAGCAAG TGCTACGAAG AGGCTAAAAA GCTGAAAAAG
 ATTAGCAACC AGGCAGAGTT CATCGCCTCC TTTTACAACA ACGACCTGAT TAAGATCAAT
 GGCGAACTGT ATAGGGTCAT CGGGGTGAAC AATGATCTGC TGAACCGCAT TGAAGTGAAT
 ATGATTGACA TCACTTACCG AGAGTATCTG GAAAACATGA ATGATAAGCG CCCCCCTCGA
 ATTATCAAAA CAATTGCCTC TAAGACTCAG AGTATCAAAA AGTACTCAAC CGACATTCTG
 GGAAACCTGT ATGAGGTGAA GAGCAAAAAG CACCCTCAGA TTATCAAAAA GGGCTAAGAA
 TTC [SEQ ID NO: 32]

[00145] SEQ ID NO: 43 is set forth below.

ATGGCCCCAAAGAAGAAGCGGAAGGTCGGTATCCACGGAGTCCCAGCAGCCAAGCGGAACCTACAT
 CCTGGGCCTGGACATCGGCATCACCAGCGTGGGCTACGGCATCATCGACTACGAGACACGGGACG
 TGATCGATGCCGGCGTGCGGCTGTTCAAAGAGGCCAACGTGGAAAACAACGAGGGCAGGCGGAGC
 AAGAGAGGCGCCAGAAGGCTGAAGCGGCGGAGGCGGCATAGAATCCAGAGAGTGAAGAAGCTGCT
 GTTCGACTACAACCTGCTGACCGACCACAGCGAGCTGAGCGGCATCAACCCCTACGAGGCCAGAG
 TGAAGGGCCTGAGCCAGAAGCTGAGCGAGGAAGAGTTCTCTGCCGCCCTGCTGCACCTGGCCAAG
 AGAAGAGGCGTGACACACGTGAACGAGGTGGAAGAGGACACCGGCAACGAGCTGTCCACCAAAGA
 GCAGATCAGCCGGAACAGCAAGGCCCTGGAAGAGAAATACGTGGCCGAACCTGCAGCTGGAACGGC
 TGAAGAAAGACGGCGAAGTGCGGGGCAGCATCAACAGATTCAAGACCAGCGACTACGTGAAAGAA
 GCCAAACAGCTGCTGAAGGTGCAGAAGGCCTACCACCAGCTGGACCAGAGCTTCATCGACACCTA
 CATCGACCTGCTGGAAACCCGGCGGACCTACTATGAGGGACCTGGCGAGGGCAGCCCCTTCGGCT
 GGAAGGACATCAAAGAATGGTACGAGATGCTGATGGGCCACTGCACCTACTTCCCCGAGGAACTG
 CGGAGCGTGAAGTACGCCTACAACGCCGACCTGTACAACGCCCTGAACGACCTGAACAATCTCGT
 GATCACCAGGGACGAGAACGAGAAGCTGGAATATTACGAGAAGTTCCAGATCATCGAGAACGTGT
 TCAAGCAGAAGAAGAAGCCACCCCTGAAGCAGATCGCCAAAGAAATCCTCGTGAACGAAGAGGAT
 ATTAAGGGCTACAGAGTGACCAGCACCGGCAAGCCCGAGTTCACCAACCTGAAGGTGTACCACGA

CATCAAGGACATTACCGCCCGGAAAGAGATTATTGAGAACGCCGAGCTGCTGGATCAGATTGCCA
AGATCCTGACCATCTACCAGAGCAGCGAGGACATCCAGGAAGAACTGACCAATCTGAACTCCGAG
CTGACCCAGGAAGAGATCGAGCAGATCTCTAATCTGAAGGGCTATACCGGCACCCACAACCTGAG
CCTGAAGGCCATCAACCTGATCCTGGACGAGCTGTGGCACACCAACGACAACCAGATCGCTATCT
TCAACCGGCTGAAGCTGGTGCCCAAGAAGGTGGACCTGTCCCAGCAGAAAGAGATCCCCACCACC
CTGGTGGACGACTTCATCCTGAGCCCCGTCGTGAAGAGAAGCTTCATCCAGAGCATCAAAGTGAT
CAACGCCATCATCAAGAAGTACGGCCTGCCCAACGACATCATTATCGAGCTGGCCCGCGAGAAGA
ACTCCAAGGACGCCCAGAAAATGATCAACGAGATGCAGAAGCGGAACCGGCAGACCAACGAGCGG
ATCGAGGAAATCATCCGGACCACCGGCAAAGAGAACGCCAAGTACCTGATCGAGAAGATCAAGCT
GCACGACATGCAGGAAGGCAAGTGCCTGTACAGCCTGGAAGCCATCCCTCTGGAAGATCTGCTGA
ACAACCCCTTCAACTATGAGGTGGACCACATCATCCCCAGAAGCGTGTCTTCGACAACAGCTTC
AACAACAAGGTGCTCGTGAAGCAGGAAGAAAACAGCAAGAAGGGCAACCGGACCCCATTCAGTA
CCTGAGCAGCAGCGACAGCAAGATCAGCTACGAAACCTTCAAGAAGCACATCCTGAATCTGGCCA
AGGGCAAGGGCAGAATCAGCAAGACCAAGAAAGAGTATCTGCTGGAAGAACGGGACATCAACAGG
TTCTCCGTGCAGAAAGACTTCATCAACCGGAACCTGGTGGATACCAGATACGCCACCAGAGGCCT
GATGAACCTGCTGCGGAGCTACTTCAGAGTGAACAACCTGGACGTGAAAGTGAAGTCCATCAATG
GCGGCTTACCAGCTTTCTGCGGCGGAAGTGGAAGTTTAAGAAAGAGCGGAACAAGGGGTACAAG
CACCACGCCGAGGACGCCCTGATCATTTGCCAACGCCGATTTTCATCTTCAAAGAGTGGAAGAACT
GGACAAGGCCAAAAAAGTGATGGAAAACCAGATGTTTCGAGGAAAAGCAGGCCGAGAGCATGCCCC
AGATCGAAACCGAGCAGGAGTACAAAGAGATCTTCATCACCCCCACCAGATCAAGCACATTAAG
GACTTCAAGGACTACAAGTACAGCCACCGGGTGGACAAGAAGCCTAATAGAGAGCTGATTAACGA
CACCCCTGTACTCCACCCGGAAGGACGACAAGGGCAACACCCCTGATCGTGAACAATCTGAACGGCC
TGTACGACAAGGACAATGACAAGCTGAAAAAGCTGATCAACAAGAGCCCCGAAAAGCTGCTGATG
TACCACCACGACCCCCAGACCTACCAGAACTGAAGCTGATTATGGAACAGTACGGCGACGAGAA
GAATCCCCTGTACAAGTACTACGAGGAAACCGGGAACCTACCTGACCAAGTACTCCAAAAAGGACA
ACGGCCCCGTGATCAAGAAGATTAAGTATTACGGCAACAACTGAACGCCCATCTGGACATCACC
GACGACTACCCCAACAGCAGAAACAAGGTCGTGAAGCTGTCCCTGAAGCCCTACAGATTGACGT
GTACCTGGACAATGGCGTGTACAAGTTCGTGACCGTGAAGAATCTGGATGTGATCAAAAAAGAAA
ACTACTACGAAGTGAATAGCAAGTGCTATGAGGAAGCTAAGAAGCTGAAGAAGATCAGCAACCAG
GCCGAGTTTATCGCCTCCTTCTACAACAACGATCTGATCAAGATCAACGGCGAGCTGTATAGAGT
GATCGGCGTGAACAACGACCTGCTGAACCGGATCGAAGTGAACATGATCGACATCACCTACCGCG
AGTACCTGGAAAACATGAACGACAAGAGGCCCCCAGGATCATTAAAGACAATCGCCTCCAAGACC
CAGAGCATTAAGAAGTACAGCACAGACATTCTGGGCAACCTGTATGAAGTGAAATCTAAGAAGCA

CCCTCAGATCATCAAAAAGGGCAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAGA
AAAAG [SEQ ID NO: 43]

[00146] In some embodiments, the nucleotide sequence encoding a *S. aureus* Cas9 molecule includes a nucleotide sequence of SEQ ID NO: 44, which is provided below:

AAGCGGAAC TACATCCTGGGCCTGGACATCGGCATCACCAGCGTGGGCTACGGCATCATCGACTA
CGAGACACGGGACGTGATCGATGCCGGCGTGCGGCTGTTCAAAGAGGCCAACGTGGAAAACAACG
AGGGCAGGCGGAGCAAGAGAGGGCGCCAGAAGGCTGAAGCGGCGGAGGCGGCATAGAATCCAGAGA
GTGAAGAAGCTGCTGTTCTGACTACAACCTGCTGACCGACCACAGCGAGCTGAGCGGCATCAACCC
CTACGAGGCCAGAGTGAAGGGCCTGAGCCAGAAGCTGAGCGAGGAAGAGTTCTCTGCCGCCCTGC
TGCACCTGGCCAAGAGAAGAGGCGTGACACAACGTGAACGAGGTGGAAGAGGACACCGGCAACGAG
CTGTCCACCAAAGAGCAGATCAGCCGGAACAGCAAGGCCCTGGAAGAGAAATACGTGGCCGAACT
GCAGCTGGAACGGCTGAAGAAAGACGGCGAAGTGCGGGGCAGCATCAACAGATTCAAGACCAGCG
ACTACGTGAAAGAAGCCAAACAGCTGCTGAAGGTGCAGAAGGCCTACCACCAGCTGGACCAGAGC
TTCATCGACACCTACATCGACCTGCTGGAACCCGGCGGACCTACTATGAGGGACCTGGCGAGGG
CAGCCCCCTTCGGCTGGAAGGACATCAAAGAATGGTACGAGATGCTGATGGGCCACTGCACCTACT
TCCCCGAGGAAC TGCGGAGCGTGAAGTACGCCTACAACGCCGACCTGTACAACGCCCTGAACGAC
CTGAACAATCTCGTGATCACCAGGGACGAGAACGAGAAGCTGGAATATTACGAGAAGTTCCAGAT
CATCGAGAACGTGTTCAAGCAGAAGAAGAAGCCCACCCTGAAGCAGATCGCCAAAGAAATCCTCG
TGAACGAAGAGGATATTAAGGGCTACAGAGTGACCAGCACCGGCAAGCCCGAGTTCACCAACCTG
AAGGTGTACCACGACATCAAGGACATTACCGCCCCGAAAGAGATTATTGAGAACGCCGAGCTGCT
GGATCAGATTGCCAAGATCCTGACCATCTACCAGAGCAGCGAGGACATCCAGGAAGAACTGACCA
ATCTGAACTCCGAGCTGACCCAGGAAGAGATCGAGCAGATCTCTAATCTGAAGGGCTATACCGGC
ACCCACAACCTGAGCCTGAAGGCCATCAACCTGATCCTGGACGAGCTGTGGCACACCAACGACAA
CCAGATCGCTATCTTCAACCGGCTGAAGCTGGTGCCCAAGAAGGTGGACCTGTCCCAGCAGAAAG
AGATCCCCACCACCCTGGTGGACGACTTCATCCTGAGCCCCGTCGTGAAGAGAAGCTTCATCCAG
AGCATCAAAGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAACGACATCATTATCGAGCT
GGCCCCGCGAGAAGAACTCCAAGGACGCCCAGAAAATGATCAACGAGATGCAGAAGCGGAACCGGC
AGACCAACGAGCGGATCGAGGAAATCATCCGGACCACCGGCAAAGAGAACGCCAAGTACCTGATC
GAGAAGATCAAGCTGCACGACATGCAGGAAGGCAAGTGCCTGTACAGCCTGGAAGCCATCCCTCT
GGAAGATCTGCTGAACAACCCCTTCAACTATGAGGTGGACCACATCATCCCCAGAAGCGTGTCCCT
TCGACAACAGCTTCAACAACAAGGTGCTCGTGAAGCAGGAAGAAAACAGCAAGAAGGGCAACCGG
ACCCCATTCAGTACCTGAGCAGCAGCGACAGCAAGATCAGCTACGAAACCTTCAAGAAGCACAT
CCTGAATCTGGCCAAGGGCAAGGGCAGAATCAGCAAGACCAAGAAAGAGTATCTGCTGGAAGAAC
GGGACATCAACAGGTTCTCCGTGCAGAAAGACTTCATCAACCGGAACCTGGTGGATACCAGATAC

GCCACCAGAGGCCTGATGAACCTGCTGCGGAGCTACTTCAGAGTGAACAACCTGGACGTGAAAGT
 GAAGTCCATCAATGGCGGCTTCACCAGCTTTCTGCGGCGGAAGTGGAAGTTTAAGAAAGAGCGGA
 ACAAGGGGTACAAGCACCACGCCGAGGACGCCCTGATCATTGCCAACGCCGATTTTCATCTTCAAA
 GAGTGGAAGAACTGGACAAGGCCAAAAAAGTGATGGAACCAGATGTTTCGAGGAAAAGCAGGC
 CGAGAGCATGCCCCGAGATCGAAACCGAGCAGGAGTACAAAGAGATCTTCATCACCCCCCACCAGA
 TCAAGCACATTAAGGACTTCAAGGACTACAAGTACAGCCACCGGGTGGACAAGAAGCCTAATAGA
 GAGCTGATTAACGACACCCTGTACTCCACCCGGAAGGACGACAAGGGCAACACCCTGATCGTGAA
 CAATCTGAACGGCCTGTACGACAAGGACAATGACAAGCTGAAAAAGCTGATCAACAAGAGCCCCG
 AAAAGCTGCTGATGTACCACCACGACCCCCAGACCTACCAGAACTGAAGCTGATTATGGAACAG
 TACGGCGACGAGAAGAATCCCCTGTACAAGTACTACGAGGAAACCGGGAACCTACCTGACCAAGTA
 CTCCAAAAAGGACAACGGCCCCGTGATCAAGAAGATTAAGTATTACGGCAACAACTGAACGCCC
 ATCTGGACATCACCGACGACTACCCCAACAGCAGAAACAAGGTCGTGAAGCTGTCCCTGAAGCCC
 TACAGATTCGACGTGTACCTGGACAATGGCGTGTACAAGTTCGTGACCGTGAAGAATCTGGATGT
 GATCAAAAAAGAAAATACTACGAAGTGAATAGCAAGTGCTATGAGGAAGCTAAGAAGCTGAAGA
 AGATCAGCAACCAGGCCGAGTTTATCGCCTCCTTCTACAACAACGATCTGATCAAGATCAACGGC
 GAGCTGTATAGAGTGATCGGCGTGAACAACGACCTGCTGAACCGGATCGAAGTGAACATGATCGA
 CATCACCTACCGCGAGTACCTGGAAAACATGAACGACAAGAGGCCCCCAGGATCATTAAGACAA
 TCGCCTCCAAGACCCAGAGCATTAAGAAGTACAGCACAGACATTCTGGGCAACCTGTATGAAGTG
 AAATCTAAGAAGCACCCCTCAGATCATCAAAAAGGGC [SEQ ID NO: 44]

[00147] An amino acid sequence of an *S. aureus* Cas9 molecule is set forth in SEQ ID NO: 33, which is provided below.

MKRNYYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQ
 RVKKLLFDYNLLTDHSELSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGN
 ELSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTSYVKEAKQLLKVQKAYHQLDQ
 SFIDTYIDLLETRRTYYEGPGEPSFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALN
 DLNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTN
 LKVYHDIKDITARKEI IENAELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNLKGYT
 GTHNLSLKAINLILDELWHTNDNQIAIFNRLKLVKKVDLSQQKEIPTTLVDDFILSPVVKRSFI
 QSIKVINAI IKKYGLPNDII IELAREKNSKDAQKMINEMQKRNRQTNERIEEII RTTGKENAKYL
 IEKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHI IPRSVSFDNSFNNKVLVKQEENSCKGN
 RTPFQYLSSSDSKI SYETFKKHILNLAKGKGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTR
 YATRGLMNNLLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHAEDALI IANADFI F
 KEWKKLDKAKKVMENQMFEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKN
 RELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLKKLINKSPEKLLMYHHPQTYQKLKLIME

QYGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNHLDDITDDYPNSRNKVVKLSLK
 PYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKIN
 GELYRVIGVNNDLLNRIEVMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYE
 VKSKKHPQIIKKG [SEQ ID NO: 33]

[00148] An amino acid sequence of an *S. aureus* Cas9 molecule is set forth in SEQ ID NO: 45, which is provided below.

KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQR
 VKKLDFDYNLLTDHSELGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNE
 LSTKEQISRNSKALEEKYVAELQLERLKKDGEVRGSINRFKTS DYVKEAKQLLKVKQKAYHQLDQS
 FIDTYIDLLETRRTYYEGPGEKSPFGWKDIKEWYEMLMGHCTYFPEELRSVKYAYNADLYNALND
 LNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTNL
 KVYHDIKDITARKEIENAEELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNLKGYTG
 THNLSLKAINLILDELWHTNDNQIAIFNRLKLVPKKVDLSQQKEIPTTLVDDFILSPVVKRSFIQ
 SIKVINAI IKKYGLPNDIIIELAREKNSKDAQKMINEMQKRNRQTNERIEEII RTTGKENAKYLI
 EKIKLHDMQEGKCLYSLEAIPLEDLLNPFNYEVDHII PRSVSFDNSFNNKVLVKQEENSKKGNR
 TPFQYLSSSDSKISYETFKKHILNLA KGGRISKTKKEYLLEERDINRFSVQKDFINRNLVDTRY
 ATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHAEDALI IANADFIFK
 EWKKLDKAKKVMENQMFEKQAESMPEIETE QEYKEIFITPHQIKHIKDFKDYKYSRVDKKPNR
 ELINDTLYSTRKDDKGNTLIVNNLNGLYDKDNDKLLKLINKSPEKLLMYHHDPQTYQKLKLIMEQ
 YGDEKNPLYKYYEETGNYLTKYSKKDNGPVIKKIKYYGNKLNHLDDITDDYPNSRNKVVKLSLK
 YRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIKIN
 ELYRVIGVNNDLLNRIEVMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYE
 VKSKKHPQIIKKG [SEQ ID NO: 45]

[00149] Alternatively or additionally, the CRISPR/Cas9-based gene editing system can include a fusion protein. The fusion protein can comprise two heterologous polypeptide domains, wherein the first polypeptide domain comprises a Cas protein and the second polypeptide domain has an activity such as transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, nuclease activity, nucleic acid association activity, methylase activity, or demethylase activity. The fusion protein can include a Cas9 protein or a mutated Cas9 protein, fused to a second polypeptide domain that has an activity such as transcription activation activity, transcription repression activity, transcription release factor activity, histone modification activity, nuclease activity, nucleic acid association activity, methylase activity, or demethylase activity.

(a) Transcription Activation Activity

[00150] The second polypeptide domain can have transcription activation activity, i.e., a transactivation domain. For example, gene expression of endogenous mammalian genes, such as human genes, can be achieved by targeting a fusion protein of iCas9 and a transactivation domain to mammalian promoters via combinations of gRNAs. The transactivation domain can include a VP 16 protein, multiple VP 16 proteins, such as a VP48 domain or VP64 domain, or p65 domain of NF kappa B transcription activator activity. For example, the fusion protein may be iCas9-VP64.

(b) Transcription Repression Activity

[00151] The second polypeptide domain can have transcription repression activity. The second polypeptide domain can have a Kruppel associated box activity, such as a KRAB domain, ERF repressor domain activity, Mxil repressor domain activity, SID4X repressor domain activity, Mad-SID repressor domain activity or TATA box binding protein activity. For example, the fusion protein may be dCas9-KRAB.

(c) Transcription Release Factor Activity

[00152] The second polypeptide domain can have transcription release factor activity. The second polypeptide domain can have eukaryotic release factor 1 (ERF1) activity or eukaryotic release factor 3 (ERF3) activity.

(d) Histone Modification Activity

[00153] The second polypeptide domain can have histone modification activity. The second polypeptide domain can have histone deacetylase, histone acetyltransferase, histone demethylase, or histone methyltransferase activity. The histone acetyltransferase may be p300 or CREB-binding protein (CBP) protein, or fragments thereof. For example, the fusion protein may be dCas9-p300.

(e) Nuclease Activity

[00154] The second polypeptide domain can have nuclease activity that is different from the nuclease activity of the Cas9 protein. A nuclease, or a protein having nuclease activity, is an enzyme capable of cleaving the phosphodiester bonds between the nucleotide subunits of nucleic acids. Nucleases are usually further divided into endonucleases and exonucleases, although

some of the enzymes may fall in both categories. Well known nucleases are deoxyribonuclease and ribonuclease.

(f) Nucleic Acid Association Activity

[00155] The second polypeptide domain can have nucleic acid association activity or nucleic acid binding protein-DNA-binding domain (DBD) is an independently folded protein domain that contains at least one motif that recognizes double- or single-stranded DNA. A DBD can recognize a specific DNA sequence (a recognition sequence) or have a general affinity to DNA. nucleic acid association region selected from the group consisting of helix-turn-helix region, leucine zipper region, winged helix region, winged helix-turn-helix region, helix-loop-helix region, immunoglobulin fold, B3 domain, Zinc finger, HMG-box, Wor3 domain, TAL effector DNA-binding domain.

(g) Methylase Activity

[00156] The second polypeptide domain can have methylase activity, which involves transferring a methyl group to DNA, RNA, protein, small molecule, cytosine or adenine. The second polypeptide domain may include a DNA methyltransferase.

(h) Demethylase Activity

[00157] The second polypeptide domain can have demethylase activity. The second polypeptide domain can include an enzyme that remove methyl (CH₃-) groups from nucleic acids, proteins (in particular histones), and other molecules. Alternatively, the second polypeptide can convert the methyl group to hydroxymethylcytosine in a mechanism for demethylating DNA. The second polypeptide can catalyze this reaction. For example, the second polypeptide that catalyzes this reaction can be Tet1.

(b) gRNA Targeting the *Dystrophin* Gene

[00158] The CRISPR/Cas9-based gene editing system includes at least one gRNA molecule, e.g., two gRNA molecules. The gRNA provides the targeting of a CRISPR/Cas9-based gene editing system. The gRNA is a fusion of two noncoding RNAs: a crRNA and a tracrRNA. The sgRNA may target any desired DNA sequence by exchanging the sequence encoding a 20 bp protospacer which confers targeting specificity through complementary base pairing with the desired DNA target. gRNA mimics the naturally occurring crRNA:tracrRNA duplex involved in the Type II Effector system. This duplex, which may include, for example, a 42-nucleotide

crRNA and a 75-nucleotide tracrRNA, acts as a guide for the Cas9 to cleave the target nucleic acid. The “target region”, “target sequence” or “protospacer” as used interchangeably herein refers to the region of the target gene (e.g., a *dystrophin* gene) to which the CRISPR/Cas9-based gene editing system targets. The CRISPR/Cas9-based gene editing system may include at least one gRNA, wherein the gRNAs target different DNA sequences. The target DNA sequences may be overlapping. The target sequence or protospacer is followed by a PAM sequence at the 3' end of the protospacer. Different Type II systems have differing PAM requirements. For example, the *Streptococcus pyogenes* Type II system uses an “NGG” sequence, where “N” can be any nucleotide. In some embodiments, the PAM sequence may be “NGG”, where “N” can be any nucleotide. In some embodiments, the PAM sequence may be NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25).

[00159] The number of gRNA molecule encoded by a presently disclosed genetic construct (e.g., an AAV vector) can be at least 1 gRNA, at least 2 different gRNA, at least 3 different gRNA at least 4 different gRNA, at least 5 different gRNA, at least 6 different gRNA, at least 7 different gRNA, at least 8 different gRNA, at least 9 different gRNA, at least 10 different gRNAs, at least 11 different gRNAs, at least 12 different gRNAs, at least 13 different gRNAs, at least 14 different gRNAs, at least 15 different gRNAs, at least 16 different gRNAs, at least 17 different gRNAs, at least 18 different gRNAs, at least 18 different gRNAs, at least 20 different gRNAs, at least 25 different gRNAs, at least 30 different gRNAs, at least 35 different gRNAs, at least 40 different gRNAs, at least 45 different gRNAs, or at least 50 different gRNAs. The number of gRNA encoded by a presently disclosed vector can be between at least 1 gRNA to at least 50 different gRNAs, at least 1 gRNA to at least 45 different gRNAs, at least 1 gRNA to at least 40 different gRNAs, at least 1 gRNA to at least 35 different gRNAs, at least 1 gRNA to at least 30 different gRNAs, at least 1 gRNA to at least 25 different gRNAs, at least 1 gRNA to at least 20 different gRNAs, at least 1 gRNA to at least 16 different gRNAs, at least 1 gRNA to at least 12 different gRNAs, at least 1 gRNA to at least 8 different gRNAs, at least 1 gRNA to at least 4 different gRNAs, at least 4 gRNAs to at least 50 different gRNAs, at least 4 different gRNAs to at least 45 different gRNAs, at least 4 different gRNAs to at least 40 different gRNAs, at least 4 different gRNAs to at least 35 different gRNAs, at least 4 different gRNAs to at least 30 different gRNAs, at least 4 different gRNAs to at least 25 different gRNAs, at least 4 different gRNAs to at least 20 different gRNAs, at least 4 different gRNAs to at least 16 different gRNAs,

at least 4 different gRNAs to at least 12 different gRNAs, at least 4 different gRNAs to at least 8 different gRNAs, at least 8 different gRNAs to at least 50 different gRNAs, at least 8 different gRNAs to at least 45 different gRNAs, at least 8 different gRNAs to at least 40 different gRNAs, at least 8 different gRNAs to at least 35 different gRNAs, 8 different gRNAs to at least 30 different gRNAs, at least 8 different gRNAs to at least 25 different gRNAs, 8 different gRNAs to at least 20 different gRNAs, at least 8 different gRNAs to at least 16 different gRNAs, or 8 different gRNAs to at least 12 different gRNAs. In certain embodiments, the genetic construct (e.g., an AAV vector) encodes one gRNA molecule, *i.e.*, a first gRNA molecule, and optionally a Cas9 molecule. In certain embodiments, a first genetic construct (e.g., a first AAV vector) encodes one gRNA molecule, *i.e.*, a first gRNA molecule, and optionally a Cas9 molecule, and a second genetic construct (e.g., a second AAV vector) encodes one gRNA molecule, *i.e.*, a second gRNA molecule, and optionally a Cas9 molecule.

[00160] The gRNA molecule comprises a targeting domain, which is a complementary polynucleotide sequence of the target DNA sequence followed by a PAM sequence. The gRNA may comprise a “G” at the 5’ end of the targeting domain or complementary polynucleotide sequence. The targeting domain of a gRNA molecule may comprise at least a 10 base pair, at least a 11 base pair, at least a 12 base pair, at least a 13 base pair, at least a 14 base pair, at least a 15 base pair, at least a 16 base pair, at least a 17 base pair, at least a 18 base pair, at least a 19 base pair, at least a 20 base pair, at least a 21 base pair, at least a 22 base pair, at least a 23 base pair, at least a 24 base pair, at least a 25 base pair, at least a 30 base pair, or at least a 35 base pair complementary polynucleotide sequence of the target DNA sequence followed by a PAM sequence. In certain embodiments, the targeting domain of a gRNA molecule has 19-25 nucleotides in length. In certain embodiments, the targeting domain of a gRNA molecule is 20 nucleotides in length. In certain embodiments, the targeting domain of a gRNA molecule is 21 nucleotides in length. In certain embodiments, the targeting domain of a gRNA molecule is 22 nucleotides in length. In certain embodiments, the targeting domain of a gRNA molecule is 23 nucleotides in length.

[00161] The gRNA may target a region of the *dystrophin* gene (*DMD*). In certain embodiments, the gRNA can target at least one of exons, introns, the promoter region, the enhancer region, the transcribed region of the *dystrophin* gene. In certain embodiments, the gRNA molecule targets intron 50 of the human *dystrophin* gene. In certain embodiments, the

gRNA molecule targets intron 51 of the human *dystrophin* gene. In certain embodiments, the gRNA molecule targets exon 51 of the human *dystrophin* gene. The gRNA may include a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, or a complement thereof.

[00162] Single or multiplexed gRNAs can be designed to restore the dystrophin reading frame by targeting the mutational hotspot at exon 51 or and introducing either intraexonic small insertions and deletions, or excision of exon 51. Following treatment with a presently disclosed vector, dystrophin expression can be restored in Duchenne patient muscle cells *in vitro*. Human dystrophin was detected *in vivo* following transplantation of genetically corrected patient cells into immunodeficient mice. Significantly, the unique multiplex gene editing capabilities of the CRISPR/Cas9-based gene editing system enable efficiently generating large deletions of this mutational hotspot region that can correct up to 62% of patient mutations by universal or patient-specific gene editing approaches. In some embodiments, candidate gRNAs are evaluated and chosen based on off-target activity, on-target activity as measured by surveyor, and distance from the exon.

3. DNA targeting compositions

[00163] The present invention is also directed to DNA targeting compositions that comprise such genetic constructs. The DNA targeting compositions include at least one gRNA molecule (e.g., two gRNA molecules) that targets a *dystrophin* gene (e.g., human *dystrophin* gene), as described above. The at least one gRNA molecule can bind and recognize a target region. The target regions can be chosen immediately upstream of possible out-of-frame stop codons such that insertions or deletions during the repair process restore the dystrophin reading frame by frame conversion. Target regions can also be splice acceptor sites or splice donor sites, such that insertions or deletions during the repair process disrupt splicing and restore the dystrophin reading frame by splice site disruption and exon exclusion. Target regions can also be aberrant stop codons such that insertions or deletions during the repair process restore the dystrophin reading frame by eliminating or disrupting the stop codon.

[00164] In certain embodiments, the presently disclosed DNA targeting composition includes a first gRNA and a second gRNA, wherein the first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, or a complement thereof. In certain embodiments, the first gRNA molecule and the second gRNA molecule comprise different targeting domains. In certain embodiments, the first gRNA molecule is SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, or SEQ ID NO: 15, and the second gRNA molecule is SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, or SEQ ID NO: 19. In certain embodiments, the first gRNA molecule is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14 and SEQ ID NO: 15, and the second gRNA molecule is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, and SEQ ID NO: 19.

[00165] In certain embodiments, the first gRNA molecule and the second gRNA molecule are selected from the group consisting of: (i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2; (ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (vi) a first gRNA molecule

comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; (xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; and (xii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 41, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 42. In some embodiments, the DNA targeting composition includes a nucleotide sequence set forth in SEQ ID NO: 37 and/or a nucleotide sequence set forth in SEQ ID NO: 38.

[00166] In certain embodiments, the DNA targeting composition may further include at least one Cas9 molecule or a Cas9 fusion protein that recognizes a PAM of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25). In some embodiments, the DNA targeting composition includes a nucleotide sequence set forth in SEQ ID NO: 83 or SEQ ID NO: 84. In certain embodiments, the vector is configured to form a first and a second double strand break in a first and a second intron flanking exon 51 of the human *dystrophin* gene, respectively, thereby deleting a segment of the *dystrophin* gene comprising exon 51.

[00167] The deletion efficiency of the presently disclosed vectors can be related to the deletion size, *i.e.*, the size of the segment deleted by the vectors. In certain embodiments, the length or size of specific deletions is determined by the distance between the PAM sequences in the gene

being targeted (e.g., a *dystrophin* gene). In certain embodiments, a specific deletion of a segment of the *dystrophin* gene, which is defined in terms of its length and a sequence it comprises (e.g., exon 51), is the result of breaks made adjacent to specific PAM sequences within the target gene (e.g., a *dystrophin* gene).

[00168] In certain embodiments, the deletion size is about 50 to about 2,000 base pairs (bp), e.g., about 50 to about 1999 bp, about 50 to about 1900 bp, about 50 to about 1800 bp, about 50 to about 1700 bp, about 50 to about 1650 bp, about 50 to about 1600 bp, about 50 to about 1500 bp, about 50 to about 1400 bp, about 50 to about 1300 bp, about 50 to about 1200 bp, about 50 to about 1150 bp, about 50 to about 1100 bp, about 50 to about 1000 bp, about 50 to about 900 bp, about 50 to about 850 bp, about 50 to about 800 bp, about 50 to about 750 bp, about 50 to about 700 bp, about 50 to about 600 bp, about 50 to about 500 bp, about 50 to about 400 bp, about 50 to about 350 bp, about 50 to about 300 bp, about 50 to about 250 bp, about 50 to about 200 bp, about 50 to about 150 bp, about 50 to about 100 bp, about 100 to about 1999 bp, about 100 to about 1900 bp, about 100 to about 1800 bp, about 100 to about 1700 bp, about 100 to about 1650 bp, about 100 to about 1600 bp, about 100 to about 1500 bp, about 100 to about 1400 bp, about 100 to about 1300 bp, about 100 to about 1200 bp, about 100 to about 1150 bp, about 100 to about 1100 bp, about 100 to about 1000 bp, about 100 to about 900 bp, about 100 to about 850 bp, about 100 to about 800 bp, about 100 to about 750 bp, about 100 to about 700 bp, about 100 to about 600 bp, about 100 to about 500 bp, about 100 to about 400 bp, about 100 to about 350 bp, about 100 to about 300 bp, about 100 to about 250 bp, about 100 to about 200 bp, about 100 to about 150 bp, about 200 to about 1999 bp, about 200 to about 1900 bp, about 200 to about 1800 bp, about 200 to about 1700 bp, about 200 to about 1650 bp, about 200 to about 1600 bp, about 200 to about 1500 bp, about 200 to about 1400 bp, about 200 to about 1300 bp, about 200 to about 1200 bp, about 200 to about 1150 bp, about 200 to about 1100 bp, about 200 to about 1000 bp, about 200 to about 900 bp, about 200 to about 850 bp, about 200 to about 800 bp, about 200 to about 750 bp, about 200 to about 700 bp, about 200 to about 600 bp, about 200 to about 500 bp, about 200 to about 400 bp, about 200 to about 350 bp, about 200 to about 300 bp, about 200 to about 250 bp, about 300 to about 1999 bp, about 300 to about 1900 bp, about 300 to about 1800 bp, about 300 to about 1700 bp, about 300 to about 1650 bp, about 300 to about 1600 bp, about 300 to about 1500 bp, about 300 to about 1400 bp, about 300 to about 1300 bp, about 300 to about 1200 bp, about 300 to about 1150 bp, about 300 to about 1100 bp, about 300 to about

1000 bp, about 300 to about 900 bp, about 300 to about 850 bp, about 300 to about 800 bp, about 300 to about 750 bp, about 300 to about 700 bp, about 300 to about 600 bp, about 300 to about 3000 bp, about 300 to about 400 bp, or about 300 to about 350 bp. In certain embodiments, the deletion size can be about 118 base pairs, about 233 base pairs, about 326 base pairs, about 766 base pairs, about 805 base pairs, or about 1611 base pairs.

4. Compositions for Genome Editing in Muscle

[00169] The present invention is directed to genetic constructs (e.g., vectors) or a composition thereof for genome editing a target gene in skeletal muscle or cardiac muscle of a subject. The composition includes a modified AAV vector and a nucleotide sequence encoding a CRISPR/Cas9-based gene editing system, e.g., a gRNA molecule and a Cas9 molecule. The composition delivers active forms of CRISPR/Cas9-based gene editing systems to skeletal muscle or cardiac muscle. The presently disclosed genetic constructs (e.g., vectors) can be used in correcting or reducing the effects of mutations in the *dystrophin* gene involved in genetic diseases and/or other skeletal or cardiac muscle conditions, e.g., DMD. The composition may further comprise a donor DNA or a transgene. These compositions may be used in genome editing, genome engineering, and correcting or reducing the effects of mutations in genes involved in genetic diseases and/or other skeletal or cardiac muscle conditions.

a. CRISPR/Cas9-based gene editing system for targeting dystrophin

[00170] A CRISPR/Cas9-based gene editing system specific for *dystrophin* gene are disclosed herein. The CRISPR/Cas9-based gene editing system may include Cas9 and at least one gRNA to target the *dystrophin* gene. The CRISPR/Cas9-based gene editing system may bind and recognize a target region. The target regions may be chosen immediately upstream of possible out-of-frame stop codons such that insertions or deletions during the repair process restore the dystrophin reading frame by frame conversion. Target regions may also be splice acceptor sites or splice donor sites, such that insertions or deletions during the repair process disrupt splicing and restore the dystrophin reading frame by splice site disruption and exon exclusion. Target regions may also be aberrant stop codons such that insertions or deletions during the repair process restore the dystrophin reading frame by eliminating or disrupting the stop codon.

[00171] The gRNA may target a nucleotide sequence selected from the group consisting of SEQ ID NO: 1-19, 41, 42, or a complement thereof. For example, the disclosed CRISPR/Cas9-based gene editing systems were engineered to mediate highly efficient gene editing at exon 51

of the *dystrophin* gene. These CRISPR/Cas9-based gene editing systems restored dystrophin protein expression in cells from DMD patients. In some embodiments, the DNA targeting composition includes a nucleotide sequence set forth in SEQ ID NO: 37, a nucleotide sequence set forth in SEQ ID NO: 38, a nucleotide sequence set forth in SEQ ID NO: 83, and/or a nucleotide sequence set forth in SEQ ID NO: 84. For example, the DNA targeting composition includes a nucleotide sequence set forth in SEQ ID NO: 37, a nucleotide sequence set forth in SEQ ID NO: 38, and a nucleotide sequence set forth in SEQ ID NO: 83, or the DNA targeting composition includes a nucleotide sequence set forth in SEQ ID NO: 37, a nucleotide sequence set forth in SEQ ID NO: 38, and a nucleotide sequence set forth in SEQ ID NO: 84.

b. Adeno-Associated Virus Vectors

[00172] The composition may also include a viral delivery system. In certain embodiments, the vector is an adeno-associated virus (AAV) vector. The AAV vector is a small virus belonging to the genus Dependovirus of the Parvoviridae family that infects humans and some other primate species. AAV vectors may be used to deliver CRISPR/Cas9-based gene editing systems using various construct configurations. For example, AAV vectors may deliver Cas9 and gRNA expression cassettes on separate vectors or on the same vector. Alternatively, if the small Cas9 proteins, derived from species such as *Staphylococcus aureus* or *Neisseria meningitidis*, are used then both the Cas9 and up to two gRNA expression cassettes may be combined in a single AAV vector within the 4.7 kb packaging limit.

[00173] In certain embodiments, the AAV vector is a modified AAV vector. The modified AAV vector may have enhanced cardiac and skeletal muscle tissue tropism. The modified AAV vector may be capable of delivering and expressing the CRISPR/Cas9-based gene editing system in the cell of a mammal. For example, the modified AAV vector may be an AAV-SASTG vector (Piacentino *et al.* (2012) Human Gene Therapy 23:635–646). The modified AAV vector may deliver nucleases to skeletal and cardiac muscle *in vivo*. The modified AAV vector may be based on one or more of several capsid types, including AAV1, AAV2, AAV5, AAV6, AAV8, and AAV9. The modified AAV vector may be based on AAV2 pseudotype with alternative muscle-tropic AAV capsids, such as AAV2/1, AAV2/6, AAV2/7, AAV2/8, AAV2/9, AAV2.5 and AAV/SASTG vectors that efficiently transduce skeletal muscle or cardiac muscle by systemic and local delivery (Seto *et al.* Current Gene Therapy (2012) 12:139-151). The modified AAV vector may be AAV2i8G9 (Shen *et al.* J. Biol. Chem. (2013) 288:28814-28823). In some

embodiments, the composition includes a nucleotide sequence set forth in SEQ ID NO: 39 and/or a nucleotide sequence set forth in SEQ ID NO: 40. In some embodiments, the composition includes a first vector comprises a nucleotide sequence set forth in SEQ ID NO: 39 and the second vector comprises a nucleotide sequence set forth in SEQ ID NO: 40.

5. Methods of Genome Editing in Muscle

[00174] The present disclosure is directed to a method of genome editing in a skeletal muscle or cardiac muscle of a subject. The method comprises administering to the skeletal muscle or cardiac muscle of the subject the composition for genome editing in skeletal muscle or cardiac muscle, as described above. The genome editing may include correcting a mutant gene or inserting a transgene. Correcting the mutant gene may include deleting, rearranging, or replacing the mutant gene. Correcting the mutant gene may include nuclease-mediated NHEJ or HDR.

6. Methods of Correcting a Mutant Gene and Treating a Subject

[00175] The presently disclosed subject matter provides for methods of correcting a mutant gene (e.g., a mutant *dystrophin* gene, e.g., a mutant human *dystrophin* gene) in a cell and treating a subject suffering from a genetic disease, such as DMD. The method can include administering to a cell or a subject a presently disclosed genetic construct (e.g., a vector) or a composition comprising thereof as described above. The method can comprises administering to the skeletal muscle or cardiac muscle of the subject the presently disclosed genetic construct (e.g., a vector) or a composition comprising thereof for genome editing in skeletal muscle or cardiac muscle, as described above. Use of presently disclosed genetic construct (e.g., a vector) or a composition comprising thereof to deliver the CRISPR/Cas9-based gene editing system to the skeletal muscle or cardiac muscle may restore the expression of a full-functional or partially-functional protein with a repair template or donor DNA, which can replace the entire gene or the region containing the mutation. The CRISPR/Cas9-based gene editing system may be used to introduce site-specific double strand breaks at targeted genomic loci. Site-specific double-strand breaks are created when the CRISPR/Cas9-based gene editing system binds to a target DNA sequences, thereby permitting cleavage of the target DNA. This DNA cleavage may stimulate the natural DNA-repair machinery, leading to one of two possible repair pathways: homology-directed repair (HDR) or the non-homologous end joining (NHEJ) pathway.

[00176] The present disclosure is directed to genome editing with a CRISPR/Cas9-based gene editing system without a repair template, which can efficiently correct the reading frame and

restore the expression of a functional protein involved in a genetic disease. The disclosed CRISPR/Cas9-based gene editing systems may involve using homology-directed repair or nuclease-mediated non-homologous end joining (NHEJ)-based correction approaches, which enable efficient correction in proliferation-limited primary cell lines that may not be amenable to homologous recombination or selection-based gene correction. This strategy integrates the rapid and robust assembly of active CRISPR/Cas9-based gene editing systems with an efficient gene editing method for the treatment of genetic diseases caused by mutations in nonessential coding regions that cause frameshifts, premature stop codons, aberrant splice donor sites or aberrant splice acceptor sites.

a. Nuclease mediated non-homologous end joining

[00177] Restoration of protein expression from an endogenous mutated gene may be through template-free NHEJ-mediated DNA repair. In contrast to a transient method targeting the target gene RNA, the correction of the target gene reading frame in the genome by a transiently expressed CRISPR/Cas9-based gene editing system may lead to permanently restored target gene expression by each modified cell and all of its progeny. In certain embodiments, NHEJ is a nuclease mediated NHEJ, which in certain embodiments, refers to NHEJ that is initiated a Cas9 molecule, cuts double stranded DNA. The method comprises administering a presently disclosed genetic construct (e.g., a vector) or a composition comprising thereof to the skeletal muscle or cardiac muscle of the subject for genome editing in skeletal muscle or cardiac muscle.

[00178] Nuclease mediated NHEJ gene correction may correct the mutated target gene and offers several potential advantages over the HDR pathway. For example, NHEJ does not require a donor template, which may cause nonspecific insertional mutagenesis. In contrast to HDR, NHEJ operates efficiently in all stages of the cell cycle and therefore may be effectively exploited in both cycling and post-mitotic cells, such as muscle fibers. This provides a robust, permanent gene restoration alternative to oligonucleotide-based exon skipping or pharmacologic forced read-through of stop codons and could theoretically require as few as one drug treatment. NHEJ-based gene correction using a CRISPR/Cas9-based gene editing system, as well as other engineered nucleases including meganucleases and zinc finger nucleases, may be combined with other existing *ex vivo* and *in vivo* platforms for cell- and gene-based therapies, in addition to the plasmid electroporation approach described here. For example, delivery of a CRISPR/Cas9-based gene editing system by mRNA-based gene transfer or as purified cell permeable proteins

could enable a DNA-free genome editing approach that would circumvent any possibility of insertional mutagenesis.

b. Homology-Directed Repair

[00179] Restoration of protein expression from an endogenous mutated gene may involve homology-directed repair. The method as described above further includes administering a donor template to the cell. The donor template may include a nucleotide sequence encoding a full-functional protein or a partially-functional protein. For example, the donor template may include a miniaturized dystrophin construct, termed minidystrophin (“minidys”), a full-functional dystrophin construct for restoring a mutant *dystrophin* gene, or a fragment of the *dystrophin* gene that after homology-directed repair leads to restoration of the mutant *dystrophin* gene.

c. Methods of Correcting a Mutant Gene and Treating a Subject Using CRISPR/Cas9

[00180] The present disclosure is also directed to genome editing with the CRISPR/Cas9-based gene editing system to restore the expression of a full-functional or partially-functional protein with a repair template or donor DNA, which can replace the entire gene or the region containing the mutation. The CRISPR/Cas9-based gene editing system may be used to introduce site-specific double strand breaks at targeted genomic loci. Site-specific double-strand breaks are created when the CRISPR/Cas9-based gene editing system binds to a target DNA sequences using the gRNA, thereby permitting cleavage of the target DNA. The CRISPR/Cas9-based gene editing system has the advantage of advanced genome editing due to their high rate of successful and efficient genetic modification. This DNA cleavage may stimulate the natural DNA-repair machinery, leading to one of two possible repair pathways: homology-directed repair (HDR) or the non-homologous end joining (NHEJ) pathway. For example, a CRISPR/Cas9-based gene editing system directed towards the *dystrophin* gene may include a gRNA having a nucleic acid sequence of any one of SEQ ID NOs: 1-19, 41, 42, or complement thereof.

[00181] The present disclosure is directed to genome editing with CRISPR/Cas9-based gene editing system without a repair template, which can efficiently correct the reading frame and restore the expression of a functional protein involved in a genetic disease. The disclosed CRISPR/Cas9-based gene editing system and methods may involve using homology-directed repair or nuclease-mediated non-homologous end joining (NHEJ)-based correction approaches, which enable efficient correction in proliferation-limited primary cell lines that may not be

amenable to homologous recombination or selection-based gene correction. This strategy integrates the rapid and robust assembly of active CRISPR/Cas9-based gene editing system with an efficient gene editing method for the treatment of genetic diseases caused by mutations in nonessential coding regions that cause frameshifts, premature stop codons, aberrant splice donor sites or aberrant splice acceptor sites.

[00182] The present disclosure provides methods of correcting a mutant gene in a cell and treating a subject suffering from a genetic disease, such as DMD. The method may include administering to a cell or subject a CRISPR/Cas9-based gene editing system, a polynucleotide or vector encoding said CRISPR/Cas9-based gene editing system, or composition of said CRISPR/Cas9-based gene editing system as described above. The method may include administering a CRISPR/Cas9-based gene editing system, such as administering a Cas9 protein or Cas9 fusion protein containing a second domain having nuclease activity, a nucleotide sequence encoding said Cas9 protein or Cas9 fusion protein, and/or at least one gRNA, wherein the gRNAs target different DNA sequences. The target DNA sequences may be overlapping. The number of gRNA administered to the cell may be at least 1 gRNA, at least 2 different gRNA, at least 3 different gRNA at least 4 different gRNA, at least 5 different gRNA, at least 6 different gRNA, at least 7 different gRNA, at least 8 different gRNA, at least 9 different gRNA, at least 10 different gRNA, at least 15 different gRNA, at least 20 different gRNA, at least 30 different gRNA, or at least 50 different gRNA, as described above. The gRNA may include a nucleic acid sequence of at least one of SEQ ID NOs: 1-19, 41, 42, or complement thereof. The method may involve homology-directed repair or non-homologous end joining.

7. Methods of Treating Disease

[00183] The present disclosure is directed to a method of treating a subject in need thereof. The method comprises administering to a tissue of a subject the presently disclosed genetic construct (e.g., a vector) or a composition comprising thereof, as described above. In certain embodiments, the method may comprises administering to the skeletal muscle or cardiac muscle of the subject the presently disclosed genetic construct (e.g., a vector) or composition comprising thereof, as described above. In certain embodiments, the method may comprises administering to a vein of the subject the presently disclosed genetic construct (e.g., a vector) or composition comprising thereof, as described above. In certain embodiments, the subject is suffering from a skeletal muscle or cardiac muscle condition causing degeneration or weakness or a genetic

disease. For example, the subject may be suffering from Duchenne muscular dystrophy, as described above.

a. Duchenne muscular dystrophy

[00184] The method, as described above, may be used for correcting the *dystrophin* gene and recovering full-functional or partially-functional protein expression of said mutated *dystrophin* gene. In some aspects and embodiments the disclosure provides a method for reducing the effects (*e.g.*, clinical symptoms/indications) of DMD in a patient. In some aspects and embodiments the disclosure provides a method for treating DMD in a patient. In some aspects and embodiments the disclosure provides a method for preventing DMD in a patient. In some aspects and embodiments the disclosure provides a method for preventing further progression of DMD in a patient.

8. Methods of Generating a Transgenic Rodent having Δ 52 hDMD

[00185] The present disclosure is directed to a method of generating a transgenic rodent embryo having a human *dystrophin* gene with an exon 52 deletion. The method includes administering to a rodent embryo the gRNA thereby deleting exon 52 of the human *dystrophin* gene, and selecting for a transgenic rodent embryo having a deletion of exon 52 of the human dystrophin gene, wherein the rodent embryo comprises a normal human *dystrophin* gene. In some embodiments, the rodent embryo is a mouse embryo. In some embodiments, the transgenic rodent embryo is heterozygous hDMD or heterozygous hDMD- Δ 52. In some embodiments, a first gRNA molecule comprising a targeting domain that includes a nucleotide sequence set forth in SEQ ID NO: 41, and a second gRNA molecule comprising a targeting domain that includes a nucleotide sequence set forth in SEQ ID NO: 42 are administered to the rodent embryo to delete exon 52 of the human dystrophin gene. In some embodiments, the method further includes administering to the rodent embryo a Cas protein comprising an amino acid sequence set forth in SEQ ID NO: 27. The present disclosure is directed to a transgenic rodent embryo that is produced by this method. The present disclosure is also directed to a transgenic rodent produced from the transgenic rodent embryo.

9. Constructs and Plasmids

[00186] The compositions, as described above, may comprise genetic constructs that encodes the CRISPR/Cas9-based gene editing system, as disclosed herein. The genetic construct, such as a plasmid, may comprise a nucleic acid that encodes the CRISPR/Cas9-based gene editing

system, such as the Cas9 protein and Cas9 fusion proteins and/or at least one of the gRNAs. The compositions, as described above, may comprise genetic constructs that encodes the modified AAV vector and a nucleic acid sequence that encodes the CRISPR/Cas9-based gene editing system, as disclosed herein. The genetic construct, such as a plasmid, may comprise a nucleic acid that encodes the CRISPR/Cas9-based gene editing system. The compositions, as described above, may comprise genetic constructs that encodes the modified lentiviral vector, as disclosed herein.

[00187] The genetic construct, such as a recombinant plasmid or recombinant viral particle, may comprise a nucleic acid that encodes the Cas9-fusion protein and at least one gRNA. In some embodiments, the genetic construct may comprise a nucleic acid that encodes the Cas9-fusion protein and at least two different gRNAs. In some embodiments, the genetic construct may comprise a nucleic acid that encodes the Cas9-fusion protein and more than two different gRNAs. In some embodiments, the genetic construct may comprise a promoter that operably linked to the nucleotide sequence encoding the at least one gRNA molecule and/or a Cas9 molecule. In some embodiments, the promoter is operably linked to the nucleotide sequence encoding a first gRNA molecule, a second gRNA molecule, and/or a Cas9 molecule. The genetic construct may be present in the cell as a functioning extrachromosomal molecule. The genetic construct may be a linear minichromosome including centromere, telomeres or plasmids or cosmids.

[00188] The genetic construct may also be part of a genome of a recombinant viral vector, including recombinant lentivirus, recombinant adenovirus, and recombinant adenovirus associated virus. The genetic construct may be part of the genetic material in attenuated live microorganisms or recombinant microbial vectors which live in cells. The genetic constructs may comprise regulatory elements for gene expression of the coding sequences of the nucleic acid. The regulatory elements may be a promoter, an enhancer, an initiation codon, a stop codon, or a polyadenylation signal.

[00189] In certain embodiments, the genetic construct is a vector. The vector can be an Adeno-associated virus (AAV) vector, which encode at least one Cas9 molecule and at least one gRNA molecule; the vector is capable of expressing the at least one Cas9 molecule and the at least gRNA molecule, in the cell of a mammal. The vector can be a plasmid. The vectors can be used for *in vivo* gene therapy. The vector may be recombinant. The vector may comprise

heterologous nucleic acid encoding the fusion protein, such as the Cas9-fusion protein or CRISPR/Cas9-based gene editing system. The vector may be a plasmid. The vector may be useful for transfecting cells with nucleic acid encoding the Cas9-fusion protein or CRISPR/Cas9-based gene editing system, which the transformed host cell is cultured and maintained under conditions wherein expression of the Cas9-fusion protein or the CRISPR/Cas9-based gene editing system takes place.

[00190] Coding sequences may be optimized for stability and high levels of expression. In some instances, codons are selected to reduce secondary structure formation of the RNA such as that formed due to intramolecular bonding.

[00191] The vector may comprise heterologous nucleic acid encoding the CRISPR/Cas9-based gene editing system and may further comprise an initiation codon, which may be upstream of the CRISPR/Cas9-based gene editing system coding sequence, and a stop codon, which may be downstream of the CRISPR/Cas9-based gene editing system coding sequence. The initiation and termination codon may be in frame with the CRISPR/Cas9-based gene editing system coding sequence. The vector may also comprise a promoter that is operably linked to the CRISPR/Cas9-based gene editing system coding sequence. The promoter that is operably linked to the CRISPR/Cas9-based gene editing system coding sequence may be a promoter from simian virus 40 (SV40), a mouse mammary tumor virus (MMTV) promoter, a human immunodeficiency virus (HIV) promoter such as the bovine immunodeficiency virus (BIV) long terminal repeat (LTR) promoter, a Moloney virus promoter, an avian leukosis virus (ALV) promoter, a cytomegalovirus (CMV) promoter such as the CMV immediate early promoter, Epstein Barr virus (EBV) promoter, a U6 promoter, such as the human U6 promoter, or a Rous sarcoma virus (RSV) promoter. The promoter may also be a promoter from a human gene such as human ubiquitin C (hUbC), human actin, human myosin, human hemoglobin, human muscle creatine, or human metallothionein. The promoter may also be a tissue specific promoter, such as a muscle or skin specific promoter, natural or synthetic. Examples of such promoters are described in US Patent Application Publication Nos. US20040175727 and US20040192593, the contents of which are incorporated herein in their entirety. Examples of muscle-specific promoters include a Spc5-12 promoter (described in US Patent Application Publication No. US 20040192593, which is incorporated by reference herein in its entirety; Hakim et al. *Mol. Ther. Methods Clin. Dev.* (2014) 1:14002; and Lai et al. *Hum Mol Genet.* (2014) 23(12): 3189–3199),

a MHCK7 promoter (described in Salva et al., Mol. Ther. (2007) 15:320-329), a CK8 promoter (described in Park et al. PLoS ONE (2015) 10(4): e0124914), and a CK8e promoter (described in Muir et al., Mol. Ther. Methods Clin. Dev. (2014) 1:14025). In some embodiments, the expression of the gRNA and/or Cas9 protein is driven by tRNAs.

[00192] Each of the polynucleotide sequences encoding the gRNA molecule and/or Cas9 molecule may each be operably linked to a promoter. The promoters that are operably linked to the gRNA molecule and/or Cas9 molecule may be the same promoter. The promoters that are operably linked to the gRNA molecule and/or Cas9 molecule may be different promoters. The promoter may be a constitutive promoter, an inducible promoter, a repressible promoter, or a regulatable promoter.

[00193] The vector may also comprise a polyadenylation signal, which may be downstream of the CRISPR/Cas9-based gene editing system. The polyadenylation signal may be a SV40 polyadenylation signal, LTR polyadenylation signal, bovine growth hormone (bGH) polyadenylation signal, human growth hormone (hGH) polyadenylation signal, or human β -globin polyadenylation signal. The SV40 polyadenylation signal may be a polyadenylation signal from a pCEP4 vector (Invitrogen, San Diego, CA).

[00194] The vector may also comprise an enhancer upstream of the CRISPR/Cas9-based gene editing system, *i.e.*, the Cas9 protein or Cas9 fusion protein coding sequence or sgRNAs, or the CRISPR/Cas9-based gene editing system. The enhancer may be necessary for DNA expression. The enhancer may be human actin, human myosin, human hemoglobin, human muscle creatine or a viral enhancer such as one from CMV, HA, RSV or EBV. Polynucleotide function enhancers are described in U.S. Patent Nos. 5,593,972, 5,962,428, and WO94/016737, the contents of each are fully incorporated by reference. The vector may also comprise a mammalian origin of replication in order to maintain the vector extrachromosomally and produce multiple copies of the vector in a cell. The vector may also comprise a regulatory sequence, which may be well suited for gene expression in a mammalian or human cell into which the vector is administered. The vector may also comprise a reporter gene, such as green fluorescent protein ("GFP") and/or a selectable marker, such as hygromycin ("Hygro").

[00195] The vector may be expression vectors or systems to produce protein by routine techniques and readily available starting materials including Sambrook *et al.*, Molecular Cloning and Laboratory Manual, Second Ed., Cold Spring Harbor (1989), which is incorporated fully by

reference. In some embodiments the vector may comprise the nucleic acid sequence encoding the CRISPR/Cas9-based gene editing system, including the nucleic acid sequence encoding the Cas9 protein or Cas9 fusion protein and the nucleic acid sequence encoding the at least one gRNA comprising the nucleic acid sequence of at least one of SEQ ID NOs: 1-19, 41, 42, or complement thereof. In some embodiments, the Cas9 protein or Cas9 fusion protein is encoded by a nucleic acid sequence of any one of SEQ ID NO: 26. In some embodiments, the vector comprises a nucleic acid sequence of SEQ ID NO: 39 or SEQ ID NO: 40.

10. Pharmaceutical Compositions

[00196] The presently disclosed subject matter provides for compositions comprising the above-described genetic constructs. The pharmaceutical compositions according to the present invention can be formulated according to the mode of administration to be used. In cases where pharmaceutical compositions are injectable pharmaceutical compositions, they are sterile, pyrogen free and particulate free. An isotonic formulation is preferably used. Generally, additives for isotonicity may include sodium chloride, dextrose, mannitol, sorbitol and lactose. In some cases, isotonic solutions such as phosphate buffered saline are preferred. Stabilizers include gelatin and albumin. In some embodiments, a vasoconstriction agent is added to the formulation.

[00197] The composition may further comprise a pharmaceutically acceptable excipient. The pharmaceutically acceptable excipient may be functional molecules as vehicles, adjuvants, carriers, or diluents. The pharmaceutically acceptable excipient may be a transfection facilitating agent, which may include surface active agents, such as immune-stimulating complexes (ISCOMS), Freund's incomplete adjuvant, LPS analog including monophosphoryl lipid A, muramyl peptides, quinone analogs, vesicles such as squalene and squalene, hyaluronic acid, lipids, liposomes, calcium ions, viral proteins, polyanions, polycations, or nanoparticles, or other known transfection facilitating agents.

[00198] The transfection facilitating agent is a polyanion, polycation, including poly-L-glutamate (LGS), or lipid. The transfection facilitating agent is poly-L-glutamate, and more preferably, the poly-L-glutamate is present in the composition for genome editing in skeletal muscle or cardiac muscle at a concentration less than 6 mg/ml. The transfection facilitating agent may also include surface active agents such as immune-stimulating complexes (ISCOMS), Freund's incomplete adjuvant, LPS analog including monophosphoryl lipid A, muramyl peptides,

quinone analogs and vesicles such as squalene and squalene, and hyaluronic acid may also be used administered in conjunction with the genetic construct. In some embodiments, the DNA vector encoding the composition may also include a transfection facilitating agent such as lipids, liposomes, including lecithin liposomes or other liposomes known in the art, as a DNA-liposome mixture (see for example International Patent Publication No. W09324640), calcium ions, viral proteins, polyanions, polycations, or nanoparticles, or other known transfection facilitating agents. Preferably, the transfection facilitating agent is a polyanion, polycation, including poly-L-glutamate (LGS), or lipid.

11. Methods of Delivery

[00199] Provided herein is a method for delivering the presently disclosed genetic construct (e.g., a vector) or a composition thereof to a cell. The delivery of the compositions may be the transfection or electroporation of the composition as a nucleic acid molecule that is expressed in the cell and delivered to the surface of the cell. The nucleic acid molecules may be electroporated using BioRad Gene Pulser Xcell or Amaxa Nucleofector IIB devices. Several different buffers may be used, including BioRad electroporation solution, Sigma phosphate-buffered saline product #D8537 (PBS), Invitrogen OptiMEM I (OM), or Amaxa Nucleofector solution V (N.V.). Transfections may include a transfection reagent, such as Lipofectamine 2000.

[00200] Upon delivery of the presently disclosed genetic construct or composition to the tissue, and thereupon the vector into the cells of the mammal, the transfected cells will express the gRNA molecule(s) and the Cas9 molecule. The genetic construct or composition may be administered to a mammal to alter gene expression or to re-engineer or alter the genome. For example, the genetic construct or composition may be administered to a mammal to correct the *dystrophin* gene in a mammal. The mammal may be human, non-human primate, cow, pig, sheep, goat, antelope, bison, water buffalo, bovids, deer, hedgehogs, elephants, llama, alpaca, mice, rats, or chicken, and preferably human, cow, pig, or chicken.

[00201] The genetic construct (e.g., a vector) encoding the gRNA molecule(s) and the Cas9 molecule can be delivered to the mammal by DNA injection (also referred to as DNA vaccination) with and without *in vivo* electroporation, liposome mediated, nanoparticle facilitated, and/or recombinant vectors. The recombinant vector can be delivered by any viral

mode. The viral mode can be recombinant lentivirus, recombinant adenovirus, and/or recombinant adeno-associated virus.

[00202] A presently disclosed genetic construct (e.g., a vector) or a composition comprising thereof can be introduced into a cell to genetically correct a *dystrophin* gene (e.g., human *dystrophin* gene). In certain embodiments, a presently disclosed genetic construct (e.g., a vector) or a composition comprising thereof is introduced into a myoblast cell from a DMD patient. In certain embodiments, the genetic construct (e.g., a vector) or a composition comprising thereof is introduced into a fibroblast cell from a DMD patient, and the genetically corrected fibroblast cell can be treated with MyoD to induce differentiation into myoblasts, which can be implanted into subjects, such as the damaged muscles of a subject to verify that the corrected dystrophin protein is functional and/or to treat the subject. The modified cells can also be stem cells, such as induced pluripotent stem cells, bone marrow-derived progenitors, skeletal muscle progenitors, human skeletal myoblasts from DMD patients, CD 133⁺ cells, mesoangioblasts, and MyoD- or Pax7- transduced cells, or other myogenic progenitor cells. For example, the CRISPR/Cas9-based gene editing system may cause neuronal or myogenic differentiation of an induced pluripotent stem cell.

12. Routes of Administration

[00203] The presently disclosed genetic constructs (e.g., vectors) or a composition comprising thereof may be administered to a subject by different routes including orally, parenterally, sublingually, transdermally, rectally, transmucosally, topically, via inhalation, via buccal administration, intrapleurally, intravenous, intraarterial, intraperitoneal, subcutaneous, intramuscular, intranasal intrathecal, and intraarticular or combinations thereof. In certain embodiments, the presently disclosed genetic construct (e.g., a vector) or a composition is administered to a subject (e.g., a subject suffering from DMD) intramuscularly, intravenously or a combination thereof. For veterinary use, the presently disclosed genetic constructs (e.g., vectors) or compositions may be administered as a suitably acceptable formulation in accordance with normal veterinary practice. The veterinarian may readily determine the dosing regimen and route of administration that is most appropriate for a particular animal. The compositions may be administered by traditional syringes, needleless injection devices, "microprojectile bombardment gone guns", or other physical methods such as electroporation ("EP"), "hydrodynamic method", or ultrasound.

[00204] The presently disclosed genetic construct (e.g., a vector) or a composition may be delivered to the mammal by several technologies including DNA injection (also referred to as DNA vaccination) with and without *in vivo* electroporation, liposome mediated, nanoparticle facilitated, recombinant vectors such as recombinant lentivirus, recombinant adenovirus, and recombinant adenovirus associated virus. The composition may be injected into the skeletal muscle or cardiac muscle. For example, the composition may be injected into the tibialis anterior muscle or tail.

[00205] In some embodiments, the presently disclosed genetic construct (e.g., a vector) or a composition thereof is administered by 1) tail vein injections (systemic) into adult mice; 2) intramuscular injections, for example, local injection into a muscle such as the TA or gastrocnemius in adult mice; 3) intraperitoneal injections into P2 mice; or 4) facial vein injection (systemic) into P2 mice.

13. Cell types

[00206] Any of these delivery methods and/or routes of administration can be utilized with a myriad of cell types, for example, those cell types currently under investigation for cell-based therapies of DMD, including, but not limited to, immortalized myoblast cells, such as wild-type and DMD patient derived lines, for example $\Delta 48-50$ DMD, DMD 6594 (del48-50), DMD 8036 (del48-50), C25C14 and DMD-7796 cell lines, primal DMD dermal fibroblasts, induced pluripotent stem cells, bone marrow-derived progenitors, skeletal muscle progenitors, human skeletal myoblasts from DMD patients, CD 133⁺ cells, mesoangioblasts, cardiomyocytes, hepatocytes, chondrocytes, mesenchymal progenitor cells, hematopoietic stem cells, smooth muscle cells, and MyoD- or Pax7-transduced cells, or other myogenic progenitor cells. Immortalization of human myogenic cells can be used for clonal derivation of genetically corrected myogenic cells. Cells can be modified *ex vivo* to isolate and expand clonal populations of immortalized DMD myoblasts that include a genetically corrected *dystrophin* gene and are free of other nuclease-introduced mutations in protein coding regions of the genome. Alternatively, transient *in vivo* delivery of CRISPR/Cas9-based systems by non-viral or non-integrating viral gene transfer, or by direct delivery of purified proteins and gRNAs containing cell-penetrating motifs may enable highly specific correction *in situ* with minimal or no risk of exogenous DNA integration.

14. Kits

[00207] Provided herein is a kit, which may be used to correct a mutated dystrophin gene. The kit comprises at least a gRNA for correcting a mutated dystrophin gene and instructions for using the CRISPR/Cas9-based gene editing system. Also provided herein is a kit, which may be used for genome editing of a dystrophin gene in skeletal muscle or cardiac muscle. The kit comprises genetic constructs (e.g., vectors) or a composition comprising thereof for genome editing in skeletal muscle or cardiac muscle, as described above, and instructions for using said composition.

[00208] Instructions included in kits may be affixed to packaging material or may be included as a package insert. While the instructions are typically written or printed materials they are not limited to such. Any medium capable of storing such instructions and communicating them to an end user is contemplated by this disclosure. Such media include, but are not limited to, electronic storage media (e.g., magnetic discs, tapes, cartridges, chips), optical media (e.g., CD ROM), and the like. As used herein, the term "instructions" may include the address of an internet site that provides the instructions.

[00209] The genetic constructs (e.g., vectors) or a composition comprising thereof for correcting a mutated dystrophin or genome editing of a dystrophin gene in skeletal muscle or cardiac muscle may include a modified AAV vector that includes a gRNA molecule(s) and a Cas9 molecule, as described above, that specifically binds and cleaves a region of the *dystrophin* gene. The CRISPR/Cas9-based gene editing system, as described above, may be included in the kit to specifically bind and target a particular region in the mutated *dystrophin* gene. The kit may further include donor DNA, a different gRNA, or a transgene, as described above.

15. Examples

[00210] It will be readily apparent to those skilled in the art that other suitable modifications and adaptations of the methods of the present disclosure described herein are readily applicable and appreciable, and may be made using suitable equivalents without departing from the scope of the present disclosure or the aspects and embodiments disclosed herein. Having now described the present disclosure in detail, the same will be more clearly understood by reference to the following examples, which are merely intended only to illustrate some aspects and embodiments of the disclosure, and should not be viewed as limiting to the scope of the

disclosure. The disclosures of all journal references, U.S. patents, and publications referred to herein are hereby incorporated by reference in their entireties.

[00211] The present invention has multiple aspects, illustrated by the following non-limiting examples.

Example 1

Targeting Human *Dystrophin* Gene

[00212] A CRISPR/Cas9-based gene editing system was used to target and delete exon 51 of the human *dystrophin* gene. The *S. aureus* Cas9 (SaCas9), which is about 1 kb smaller than *S. pyogenes* Cas9, was used with an adeno-associated virus (AAV) to deliver the CRISPR/Cas9-based gene editing system. The codon optimized nucleic acid sequence encoding the *S. aureus* Cas9 molecule is set forth in SEQ ID NO: 43 or SEQ ID NO: 44. FIG. 3 shows a schematic of the AAV-based *in vivo* co-delivery of SaCas9 and two gRNAs on two viral vectors to muscle tissues. Each vector had a copy of SaCas9 and one gRNA driven by the CMV and hU6 promoters, respectively (PT366-179 (SEQ ID NO: 39) and PT366-183 (SEQ ID NO: 40)).

[00213] The activity of individual gRNAs that target the human *dystrophin* gene, JCR89 (targets upstream of exon 51) and JCR91 (targets downstream of exon 51), which were designed against the human genome, was determined by Surveyor Assay in HEK293T cells (has normal version of *dystrophin* gene) and DMD patient myoblast lines (DMD 8036 and DMD 6594, which each have mutant form of *dystrophin* gene) (see FIG. 1). The Surveyor assay detects mismatches in the genomic DNA, which is indicative of indels from the CRISPR/Cas9-based gene editing system. For JCR89, the parent band size was 555 nt and the primer used were: forward primer - aagttactgtgccagcatga (SEQ ID NO: 91); and reverse primer - gaaaaacttctgccaaactttatca (SEQ ID NO: 92). The expected cut band sizes were 134 nt and 421 nt. For JCR91, the parent band size was 632 nt and the primer used were: forward primer - tgcaaataacaaaagtagccatata (SEQ ID NO: 93); and reverse primer - tcttagaaaggcttgaaagctg (SEQ ID NO: 94). The expected cut band sizes were 210 nt and 422 nt.

[00214] HEK293T cells and DMD myoblasts (DMD 8036 and DMD 6594) were co-treated with SaCas9 and gRNAs JCR89 and JCR91 (SEQ ID NO: 37 and SEQ ID NO: 38). Genomic DNA was amplified with forward primer - cttcactgtggccagttta (SEQ ID NO: 95); and reverse primer - tcttagaaaggcttgaaagctg (SEQ ID NO: 94). The expected parent band size was 1646 nt

and the expected “perfect” deletion band was 766 nt (the actual deletion size between the gRNA cut sites varied from the 766 nt due to the occurrence of indels). FIG. 2A shows the deletion of exon 51 in genomic DNA of the HEK293T cells and DMD myoblasts. FIG. 2B shows the deletion of exon 51 in cDNA from DMD myoblasts. “No RT” is a negative control where no reverse transcriptase was added.

[00215] The CRISPR/Cas9-based gene editing system was injected into transgenic mice carrying the human DMD gene (hDMD/*mdx* mice) to delete exon 51. Local AAV8 delivery of the viral vectors carrying the SaCas9 and gRNAs was applied to tibialis anterior (TA) muscle. See Table 1. 3 mice were injected with AAV8: 1 mouse was injected with a high dose of AAV8 in both TAs (“HH”), 1 mouse was injected with low dose of AAV8 in both TAs (“LL”), and 1 mouse was injected with the low dose in the left TA and the high dose in the right TA (“LH”). Doses are listed in column 2 of Table 1. The mice were sacrificed 8 weeks post treatment (“Weeks PT”) to harvest tissues for analysis. Nested PCR revealed deletion of exon 51 in both of the limbs in the HH mouse, in the right TA of the LL mouse, and in the right limb of the LH mouse.

Table 1

| Experiment | AAV8 dosage | Delivery | Weeks PT | Nested PCR - gDNA Δ51 |
|-------------------|------------------------|---|-----------------|--|
| HH, LL, LH | HH: 6.6E11 LL: 1E11 | Intramuscular tibialis anterior (TA) | 8 | HH: both, LL: R, LH: R |

[00216] Genomic DNA harvested from mouse TA muscle was amplified in a first PCR reaction using forward primer: cttcactgctggccagtta (SEQ ID NO: 95) and reverse primer: tcttagaaaggcttgaaagctg (SEQ ID NO: 94). 1-3 μ L of the PCR product was used in a second PCR reaction (2X gDNA PCR) using forward primer – aagttactgtccaggcatga (SEQ ID NO: 91); and reverse primer – ttgaacatggcattgcataaA (SEQ ID NO: 96). This second PCR had an expected parent band of 1089 nt and an expected deletion band of 323 nt (the actual deletion size between the gRNA cut sites varied from the 323 nt due to the occurrence of indels). FIG. 4 shows the second PCR results. The “L” lanes show the results for the left TA muscle, which was used as a control and received saline solution. The “R” lanes show the results for the right TA muscle, which were injected with the 2 viral vectors that were pre-mixed in equal amounts. The CRISPR/Cas9-based gene editing system was also injected into the tail vein of hDMD/*mdx* mice

via systemic AAV8 delivery (see FIG. 5). Genomic DNA harvested from mouse liver (FIG. 5 - left panel) and heart (FIG. 5 - right panel) was also amplified using the same protocol as for FIG. 4. The expected band of approximately 300 nucleotides indicated the deletion of exon 51.

[00217] Further gRNAs that target human and rhesus monkey *dystrophin* gene sequences or human *dystrophin* gene sequences were generated and selected (see FIG. 6 and Table 2). Table 2 lists the general target of the gRNAs, the genomic strand that is recognized, the gRNA sequence, and the PAM sequence associated with the gRNA. The target genome sequences for the gRNAs, as indicated on the genomic plus strand, are listed in Table 3. These gRNAs were tested in cultured human cells to find optimal activity and combinations of gRNAs to generate deletions. Selected gRNAs were also prioritized based on predicted specificity in the human genome (FIG. 8) and screened for optimal target sequence lengths varying from 19-23 nucleotides (FIGS. 10 and 11). FIG. 6 shows the various gRNA targets listed in Table 2 that are conserved between human and rhesus macaque genomes. The location of each gRNA is indicated in relation to exon 51 of the human *dystrophin* gene.

Table 2 List of gRNAs

| gRNA | General target | Strand | gRNA sequence (5'-3') | SEQ ID NO: | PAM Sequence | SEQ ID NO: |
|--------|---|--------|------------------------|------------|--------------|------------|
| JCR89 | Human DMD exon 51 (upstream) | plus | AAAGATATATAATGTCATGAAT | 1 | AAGAGT | 53 |
| JCR91 | Human DMD exon 51 (downstream) | plus | GCAGAATCAAATATAATAGTCT | 2 | GGAAT | 54 |
| JCR94 | Human DMD exon 52 (upstream) | minus | AACAAATATCCCTTAGTATC | 41 | AGG | 55 |
| JCR99 | Human DMD exon 52 (downstream) | minus | AATGTATTTCTTCTATTCAA | 42 | TGG | 56 |
| JCR159 | Human and rhesus DMD exon 51 (upstream) | minus | AACAATAAGTCAAATTTAATTG | 3 | AAGAGT | 53 |
| JCR160 | Human and rhesus DMD exon 51 (downstream) | minus | GAACTGGTGGGAAATGGTCTAG | 4 | GAGAGT | 57 |
| JCR167 | Human and rhesus DMD exon 51 | minus | TCCTTTGGTAAATAAAAGTCCT | 5 | GGGAGT | 58 |

| | | | | | | |
|--------|---|-------|-------------------------|----|--------|----|
| | (downstream) | | | | | |
| JCR166 | Human and rhesus DMD exon 51 (downstream) | minus | TAGGAATCAAATGGACTTGGAT | 6 | TTGAAT | 59 |
| JCR168 | Human and rhesus DMD exon 51 (upstream) | plus | TAATTCTTTCTAGAAAGAGCCT | 7 | CAGAGT | 60 |
| JCR170 | Human and rhesus DMD exon 51 (upstream) | minus | CTCTTGCATCTTGCACATGTCC | 8 | TGGAGT | 61 |
| JCR171 | Human and rhesus DMD exon 51 (upstream) | minus | ACTTAGAGGTCTTCTACATACA | 9 | ATGAGT | 62 |
| JCR156 | Human and rhesus DMD exon 51 (upstream) | minus | TCAGAGGTGAGTGGTGAGGGGA | 10 | AGGAAT | 63 |
| JCR157 | Human and rhesus DMD exon 51 (upstream) | minus | ACACACAGCTGGGTTATCAGAG | 11 | GAGAGT | 57 |
| JCR176 | Human and rhesus DMD exon 51 (upstream) | minus | CACAGCTGGGTTATCAGAG | 12 | GAGAGT | 57 |
| JCR177 | Human and rhesus DMD exon 51 (upstream) | minus | ACACAGCTGGGTTATCAGAG | 13 | GAGAGT | 57 |
| JCR178 | Human and rhesus DMD exon 51 (upstream) | minus | CACACAGCTGGGTTATCAGAG | 14 | GAGAGT | 57 |
| JCR179 | Human and rhesus DMD exon 51 (upstream) | minus | AACACACAGCTGGGTTATCAGAG | 15 | GAGAGT | 57 |
| JCR180 | Human and rhesus DMD exon 51 (downstream) | minus | CTGGTGGGAAATGGTCTAG | 16 | GAGAGT | 57 |
| JCR181 | Human and rhesus DMD exon 51 (downstream) | minus | ACTGGTGGGAAATGGTCTAG | 17 | GAGAGT | 57 |
| JCR182 | Human and | minus | AACTGGTGGGAAATGGTCTAG | 18 | GAGAGT | 57 |

| | | | | | | |
|--------|--|-------|-------------------------|----|--------|----|
| | rhesus DMD exon 51 (downstream) | | | | | |
| JCR183 | Human and rhesus DMD exon 51 (downstream) | minus | AGAACTGGTGGGAAATGGTCTAG | 19 | GAGAGT | 57 |

Table 3 Target sequences of gRNAs

| gRNA Name | Target Sequence (Plus Strand 5'-3') | SEQ ID NO: | Comments |
|--------------|-------------------------------------|---------------|----------------------------|
| JCR89 | AAAGATATATAATGTCATGAAT | 64 | |
| JCR91 | GCAGAATCAAATATAATAGTCT | 65 | |
| JCR159 | CAATTAAATTTGACTTATTGTT | 66 | 101 upstream of exon 51 |
| JCR160 | CTAGACCATTTCACCAGTTC | 67 | 78 downstream of exon 51 |
| JCR167 | AGGACTTTTATTTACCAAAGGA | 68 | 1534 downstream of exon 51 |
| JCR166 | ATCCAAGTCCATTTGATTCTTA | 69 | 1266 downstream of exon 51 |
| JCR168 | TAATTCTTTCTAGAAAGAGCCT | 70 | 1824 upstream of exon 51 |
| JCR170 | GGACATGTGCAAGATGCAAGAG | 71 | 2851 upstream of exon 51 |
| JCR171 | TGTATGTAGAAGACCTCTAAGT | 72 | 2947 upstream of exon 51 |
| JCR156 | TCCCCTCACCCTCACCTCTGA | 73 | 1300 upstream of exon 51 |
| JCR157 | CTCTGATAACCCAGCTGTGTGT | 74 | 1284 upstream of exon 51 |
| JCR176 | CTCTGATAACCCAGCTGTG | 75 | JCR157 – 19 nucleotide |
| JCR177 | CTCTGATAACCCAGCTGTGT | 76 | JCR157 – 20 nucleotide |
| JCR178 | CTCTGATAACCCAGCTGTGTG | 77 | JCR157 – 21 nucleotide |
| JCR179 | CTCTGATAACCCAGCTGTGTGTT | 78 | JCR157 – 23 nucleotide |
| JCR180 | CTAGACCATTTCACCAG | 79 | JCR160 – 19 nucleotide |
| JCR181 | CTAGACCATTTCACCAGT | 80 | JCR160 – 20 nucleotide |
| JCR182 | CTAGACCATTTCACCAGTT | 81 | JCR160 – 21 nucleotide |
| JCR183 | CTAGACCATTTCACCAGTTCT | 82 | JCR160 – 23 nucleotide |

[00218] Human HEK293T cells were transfected with the individual candidate gRNAs listed in Table 2. The activity of the candidate gRNAs was determined by the Surveyor assay (see FIG. 7). For JCR160, the parent band size was 483 nt and the primer used were: forward primer - cgggcttggacagaacttac (SEQ ID NO: 97); and reverse primer - ctgcgtagtgccaaaacaaa (SEQ ID NO: 98). The expected cut band sizes were 192 nt and 291 nt. For JCR157, the parent band size was 631 nt and the primer used were: forward primer - gagatgtcttttcagctttcc (SEQ ID NO: 99); and reverse primer - gggaccttggttaagccaca (SEQ ID NO: 100). The expected cut band sizes were 147 nt and 484 nt.

[00219] The specificity of the candidate gRNAs was predicted using CasOFFinder program (Bae et al. (2014) *Bioinformatics* 30: 1473-1475; see FIG. 8). Candidate gRNAs were evaluated and chosen based on off-target activity, on-target activity as measured by Surveyor assay, and distance from the exon. The gRNAs JCR157 and JCR160 had low predicted off-target binding and were used for further testing.

[00220] HEK293T cells were transfected with and DMD 6594 cells were electroporated with a modified pDO240 plasmid containing the gRNA JCR157, a modified pDO240 plasmid containing the gRNA JCR160, and a plasmid containing SaCas9 (pDO242; SEQ ID NO:83). The parent band was predicted to be 2451 nt and the deletion band is predicted to be about 840-850 nt. FIG. 9 shows the deletion of exon 51 as determined by PCR of genomic DNA (approximately 850 nucleotides) using forward primer - tgcctttcaatcattgtttcg (SEQ ID NO: 101) and reverse primer - agaaggcaaattggcacaga (SEQ ID NO: 102). The deletion created between the gRNA cut sites was approximately 1611 nt.

[00221] FIG. 10 shows the activity of various target lengths of gRNA JCR157 (19, 20, 21, 22, and 23 nucleotides) determined by the Surveyor Assay in HEK293T cells using the primers and PCR conditions used for JCR157 in FIG. 7. FIG. 11 shows the activity of various target lengths of gRNA JCR160 (19, 20, 21, 22, and 23 nucleotides) determined by the Surveyor Assay in HEK293T cells using forward primer - cgggcttgacagaacttac (SEQ ID NO: 97); and reverse primer - ctgcgtagtgccaaaacaaa (SEQ ID NO: 98). The parent band size was predicted to be 483 nt and expected cut band sizes were 209 nt and 274 nt.

[00222] Combinations of the various target lengths of gRNA JCR157 and JCR160 (21, 22, or 23 nucleotides) were used in HEK293T cells using the conditions used in FIG. 9. FIG. 12 shows PCR of genomic DNA. The combination of JCR157 and JCR160 each having 23 nucleotide targets, had almost 50% deletion.

[00223] Each gRNA flanking exon 51 (upstream JCR179 and downstream JCR183) was performed individually with SaCas9 using a target sequences of 23 nt to demonstrate on-target nuclease activity in HEK293T cells ("293s") and DMD6594 cells ("DMD6594s") (see FIG. 13). For JCR179, the parent band size was 594 nt and the primer used were: forward primer - tgcctttcaatcattgtttcg (SEQ ID NO: 101); and reverse primer - aaggcccaaatgtgaaat (SEQ ID NO: 103). The expected cut band sizes were 594 nt and 130 nt. For JCR183, the parent band size was 731 nt and the primer used were: forward primer - gagtttgctcaaattgttactctt (SEQ ID NO: 104);

and reverse primer - ctgcgtagtgccaaaacaaa (SEQ ID NO: 98). The expected cut band sizes were 440 nt and 291 nt. FIG. 13 shows *in vitro* on-target nuclease activity by Surveyor assay.

[00224] Human HEK293T cells were transfected with and DMD myoblasts (DMD 6594s) were electroporated with plasmids containing SaCas9 and gRNAs JCR179 and JCR183 (23 nt targets of JCR157 and JCR160; SEQ ID NO: 37 and SEQ ID NO: 38). The DMD 6594 cells are immortalized DMD patient myoblasts that are already lacking exons 48-50. The parent band was predicted to be 2451 nt and the deletion band is predicted to be about 823 nt. FIG. 14 shows the *in vitro* deletion of exon 51 in genomic DNA in human HEK293T cells (left panels) and DMD 6594s cells (right panels) as determined by PCR of genomic DNA using forward primer - tgcctttcaatcattgtttcg (SEQ ID NO: 101) and reverse primer - agaaggcaaattggcacaga (SEQ ID NO: 102). The deletion created between the gRNA cut sites was approximately 1628 nt. The top panels show a schematic of the target gene of the upstream and downstream gRNAs in the HEK293T cells and DMD 6594 cells, wherein the purple indicates normally processed exons and yellow indicates mutant exons. The middle panels show the result of PCR across the genomic deletion region, wherein the asterisk indicates the deletion. The bottom panels show the droplet digital PCR of genomic DNA. In the HEK293T cells, the gRNAs and SaCas9 had 16% deletion, wherein the DMD 6594 cells had editing of about 10%.

[00225] To determine if the changes in genomic DNA were transcribed to the RNA, RNA was harvested from DMD myoblasts that were co-transfected with SaCas9 and both gRNAs and differentiated for 7 days (see FIG. 15). The RNA was reverse transcribed to cDNA and the cDNA was PCR amplified using standard methods known in the art. In FIG. 15, the bottom left panel shows PCR amplification from exon 44 to exon 52 using forward primer - tggcggcggttttcattat (SEQ ID NO: 105) and reverse primer - TTCGATCCGTAATGATTGTTCTAGCC (SEQ ID NO: 106). The parent band was predicted to be 948 nt and the deletion band is predicted to be about 715 nt. FIG. 15 shows a deletion band only in cells treated with SaCas9 and both gRNAs. The bottom right panel shows ddPCR revealing editing of about 14% of the cDNA. The *in vitro* exon 47 to 52 junction in cDNA of DMD patient myoblasts was sequenced (see FIG. 16). In FIG. 16, the sequence of the bands from the untreated cells (Control cells Δ 48-50) indicated that exons 47 to 51 were joined as expected, while the deletion band in the treated cells (Δ 48-50 + Δ 51) showed the junction of

exon 47 to 52. Thus a distinct lack of exon 51 and the disclosed system aimed at the genomic DNA level was being carried through transcription.

Example 2

Generation of $\Delta 52/mdx$ mouse

[00226] FIG. 17 shows the design for the $\Delta 52/mdx$ mouse. The hDMD/*mdx* mouse was obtained from Leiden University and manipulated to generate a relevant model for DMD, in which exon 52 is removed and the deletion results in an out-of reading frame shift and DMD genotype. The hDMD/*mdx* mouse contains a full length, wild-type human *dystrophin* gene on chromosome 5 in the *mdx* background such that no mouse dystrophin is expressed.

[00227] To generate a relevant model for DMD, the SpCas9 CRISPR/Cas9 editing system and gRNAs were used to target and delete exon 52 of the human *dystrophin* gene. Various gRNAs targeting upstream and downstream of exon 52 were tested and validated using Surveyor assay (see FIG. 18). For the upstream gRNAs, the forward primer - ctccggaatgtctccatttg (SEQ ID NO: 87) and reverse primer - TTGTGTGTCCCATGCTTGTT (SEQ ID NO: 107) were used and the parent band size was 402 nt. For JCR94 (AACAAATATCCCTTAGTATC (SEQ ID NO: 41)), the expected cut sizes were 243 nt and 159 nt. For the downstream gRNAs, the forward primer - CAACGCTGAAGAACCCTGAT (SEQ ID NO: 108) and reverse - atgaggagagactggcatc (SEQ ID NO: 88) were used and the parent band size was 509 nt. For JCR99 (AATGTATTTCTTCTATTCAA (SEQ ID NO: 42)), the expected cut sizes were 346 nt and 163 nt. Pairs of gRNAs were tested and validated by detecting the deletion of exon 51 in genomic DNA of HEK293T cells, including JCR94 and JCR99, using forward primer - ctccggaatgtctccatttg (SEQ ID NO: 87) and reverse primer - atgaggagagactggcatc (SEQ ID NO: 88) (see FIG. 19). The parent band size was 718 nt, the deletion band was 392 nt, and the deletion between gRNAs was 326 nt. The pair of JCR94 and JCR99 was used in the genome editing system to generate the $\Delta 52/mdx$ mouse. Specifically, the mouse was created by injecting JCR94 gRNA, JCR99 gRNA, and SaCas9 mRNA into mice embryos.

[00228] FIG. 20 shows DNA microinjection protocol which includes BAC recombineering service. Day 1: pregnant mares were intraperitoneally treated with serum gonadotropin to induce ovulation. Day 3: pregnant mares were intraperitoneally treated with human chorionic gonadotropin. In the 1st round, 471 embryos were produced, but only 5 plugs were visualized.

150 fertilized embryos were used (14 females superovulated). Pronuclear injections were performed with less than 50ng Cas9, and 20ng each guide. FIG. 21 shows the mouse breeding protocol to generate the transgenic mice.

[00229] The founder mice were genotyped using the following genotyping protocol. Genomic DNA (gDNA) was extracted from the tail snips of the mice using a DNEasy Blood and Tissue kit (Qiagen). To genotype each pup, gDNA was amplified using AccuPrime HiFi Taq kit as follows: i. 100 ng gDNA; ii. 2.5 μ L AccuPrime Buffer II; iii. 0.1 μ L AccuPrime HiFi Taq; iv. 1 μ L JRH261 (ctccggaatgtctccatttg (SEQ ID NO: 87)) (10 μ M); v. 1 μ L JRH264 (atgagggagagactggcatc (SEQ ID NO: 88)) (10 μ M); and vi. water up to 25 μ L total volume. The reactions were run on a thermocycler as follows: i. 95 degrees for 4 minutes; ii. 95 degrees for 30 sec; iii. 52 degrees for 30 sec; iv. 68 degrees for 1:00 min; v. Cycle steps ii—iv 35 times; and 4 degrees forever. The PCR reactions were separated on a gel (FIG. 22). The expected band sizes were 718 nt if no deletion was present (i.e., exon 52 was still present) and approximately 392 nt if there was a deletion of exon 52. As shown in FIG. 22, the founder mice 7, 63, and 76 had the exon 52 deletion.

[00230] The amplified bands were sequenced using the JRH264 primer (see FIG. 23) to sequence the rejoined ends of the targeted regions. FIG. 23 shows the sequenced region, where the bolded, underlined, and normal letters indicate native sequences and the italicized letters indicate insertions or deletions. In the expected sequence (“delta 52”), the bold letters are ligated to the underlined letters. In the founder mice, there were insertions (italicized letters) and deletions (hyphens) in this region.

[00231] The male founder mice were mated with *mdx/mdx* female to breed out chimera (FIG. 24). The litters generated from founder male 76 or founder male 63 with *mdx/mdx* female were screened and genotyped for the exon 52 deletion using the conditions used in FIG. 22 (FIG. 25 and FIG. 26, respectively). If exon 52 was deleted, then the expected band size was about 392 nt. If exon 52 was present, then the expected band size was about 718 nt. The pups 54497 and 54498 (from founder male 63 + *mdx/mdx* female breeding pair) had the exon 52 deletion and were sequenced (FIG. 27). Pups 54497 and 54498 had 92.86% identity with each other in a 392 bp sequencing read and the indels were identical.

[00232] The expression of dystrophin in healthy hDMD/*mdx* mouse and the Δ 52/*mdx* mouse was compared using fluorescent immunohistochemistry. As shown in FIG. 28, Δ 52/*mdx* mice

pups 54497 and 54498 lacked dystrophin protein. For the heart staining, exposure for laminin probe was 100ms, while exposure for dystrophin was 900ms. For the TA muscle sample, exposure for the laminin and dystrophin probes was 2.0 s. See also FIG. 29 which shows that the $\Delta 52/mdx$ mouse lacked dystrophin protein. In both the heart and TA muscle, dystrophin expression was lost in the $\Delta 52/mdx$ mouse. In the TA muscle, there were a few spontaneous revertant fibers (random splice events, or somatic mutations), but this was consistent with the *mdx* mouse model. Western blotting also indicated that the $\Delta 52/mdx$ mouse lacked dystrophin protein which was consistent with the DMD genotype, while the healthy hDMD/*mdx* mouse expresses dystrophin. See FIG. 30.

[00233] The $\Delta 52/mdx$ mouse showed similar levels of activity to the *mdx* mice after the first five minutes of an open field test. In the open field test, the mice were allowed to freely explore an open field arena (20 x 20 x 30 cm) for 30 minutes. The activity and location of the animal was automatically monitored using infrared diodes (x, y, and z axis) interfaced to a computer running Fusion Activity software (version 5.3, Omnitech, Columbus, OH). FIG. 31 shows the overall activity of the $\Delta 52/mdx$ mouse compared to *mdx* mice and hDMD/*mdx* mice as indicated by locomotion and exploration. The distance traveled and the upright vertical activities are shown in the left and right panels, respectively.

Example 3

Dystrophin Restoration by Removal of Exon 51

[00234] The removal of exon 51 can generate a Becker Muscular Dystrophy (BMD)-like genotype in the $\Delta 52/mdx$ mouse, and in theory restore dystrophin expression. The $\Delta 52/mdx$ mouse was used to demonstrate the restoration of dystrophin expression by the removal of exon 51 using the disclosed CRISPR/Cas9-based gene editing system. FIG. 32 shows the corrections strategy using SaCas9 and gRNAs to skip exon 51 by targeting the gRNAs upstream and downstream of exon 51 in the intronic region.

[00235] Standard: plasmids containing gRNAs JCR179 (upstream) or JCR183 (downstream) (SEQ ID NO: 37 and SEQ ID NO: 38, respectively) and SaCas9 were electroporated into DMD patient myoblasts. Protein was harvested from differentiated cells and analyzed using a Western blot with a dystrophin antibody. FIG. 33 shows that the genomic DNA can be edited to restore

the dystrophin protein as cells treated with all 3 components (i.e., both gRNAs and SaCas9) showed dystrophin expression.

[00236] The system was then tested with the $\Delta 52/mdx$ mouse. FIG. 34 shows the experimental design to treat $\Delta 52/mdx$ mouse using the gRNAs and SaCas9 system, including schematics of 2 viral vectors used in the experimental design. AAV8 recombinant viral constructs were created using vectors PT366-179 (SEQ ID NO: 39) and PT366-183 (SEQ ID NO: 40) and methods known in the art to produce viral particles. These viral vectors (AAV8) were co-delivered *in vivo* as two viral particles. Each viral particle contained SaCas9 and one of the gRNAs (see FIG. 3). $\Delta 52/mdx$ mice were treated with 5E11 of AAV8 recombinant viral constructs. The virus was injected intramuscularly into the right TA muscle, while the left TA muscle served as a contralateral control and was injected with PBS. After treatment, both the left and right TA muscles were removed and sections of each were taken for genomic DNA analyses. As shown in FIG. 35, PCR was performed across the region of interest and the deletion bands were noted in the treated right TA muscle on the left gel, indicating some level of gene editing. The deletion band was sequenced and the dominant product was the expected ligation 3 base pairs in from the PAM of each gRNA. FIG. 35 shows *in vivo* exon 51 deletion in right TA muscle.

[00237] Similarly, sections of the treated right TA and control left TA muscles were analyzed to determine if editing carried through to the RNA. PCR was performed across exons 50 to 53 in the cDNA. As shown in the right gel of FIG. 36, there is a lack of exon 51 and 52 in two of the treated samples. The deletion band was sequenced and the dominant product was the ligation of exon 50 to 53, as expected given the mouse already lacked exon 52 and the CRISPR/Cas9-based gene editing system removed exon 51 (see bottom right sequence chromatogram). FIG. 36 shows *in vivo* exon 51 deletion in right TA muscle.

[00238] FIG. 37 shows representative fluorescent immunohistochemical staining indicating that little dystrophin was present in the control PBS injected left TA muscle of the $\Delta 52/mdx$ mouse. Some degree of dystrophin staining in green on the control left TA may be due to revertant fibers or dead cells, which sometimes also stain green. There is a clear increase in the green dystrophin staining in the treated right TA muscle, as shown in the right photo. FIG. 37 shows *in vivo* dystrophin protein restoration in treated TA muscle.

[00239] Protein was extracted from the left and right TA muscles from 3 test mice and Western blot analysis was performed. FIG. 38 shows *in vivo* dystrophin protein restoration in treated TA

muscle. No protein expression was seen in the control left TA muscles, while all three right TA muscles displayed varying levels of dystrophin protein expression. The protein expression from the right TA of mouse 1 was the strongest, while mouse 2 and 3 had faint but nonetheless present bands. The disclosed CRISPR/Cas9-based gene editing system worked in vivo to restore dystrophin protein expression to some degree in the $\Delta 52/mdx$ mouse.

[00240] Mouse Physiology Testing. Treated mice (all male hDMD- $\Delta 52(het)/mdx(hemi)$ mice) were treated with SaCas9- and gRNA- containing AAV8 (n=10) or AAV9 (n=10) recombinant viral constructs and compared with untreated mice (hDMD - $\Delta 52/mdx$) (n=10.). The AAV recombinant viral constructs were created using vectors PT366-179 (SEQ ID NO: 39) and PT366-183 (SEQ ID NO: 40) and using methods known in the art. The treated mice had 200 μ L of virus injected into the tail vein between 6 and 8 weeks old. Mice were tested 8 weeks later.

[00241] Open field distance test. Mice were allowed free exploration of an open field arena for 30 minutes. The activity and location of animals was automatically monitored with infrared diodes interfaced to a computer running Fusion Activity software. Data was collected continuously and binned into 5 minute intervals. FIG. 39 shows the average of all time points for total distance moved in 16 week old mice that were treated at 8 weeks old compared to 16 week old mice that were not treated. The average of all time points for total rearing postures after 16 weeks is shown in FIG. 40. Statistics were: one way ANOVA, compared each column mean to the untreated mean, and Dunnett post hoc (Mean \pm SEM). The AAV8 and AAV9 treated mice show statistically significant more distance traveled than untreated age matched mice (statistically significant). All treated mice show statistically significant increased amounts of rearing postures compared to untreated age matched control.

[00242] Grip strength. The grip strength of 16 week untreated and treated mice were tested. Mice were given 3-5 trials each for the front and rear feet to test grip strength. The average trial is shown in FIG. 41. Grip strength is reported as grams force. As shown in FIG. 41, AAV9 treated mice showed statistically significant increased grip strength force in the front paws compared to untreated age matched mice. Statistics were: Two-way ANOVA, Tukey's test post hoc.

[00243] cDNA PCR. Tissues from the hearts of mice were processed using the RNEasy Plus Universal mini kit (Qiagen). The resulting RNA was reversed transcribed to cDNA using SuperScript VILO cDNA synthesis kit. 1 μ L of cDNA was PCR amplified using AccuPrime

DNA Polymerase and primers in exon 48 (forward primer: gttccagagctttacctgagaa (SEQ ID NO: 89)) and exon 54 (reverse primer: CTTTTATGAATGCTTCTCCAAG (SEQ ID NO: 90)). The expected band sizes were 997 nt if no deletion was present (i.e., exon 52 was still present) and approximately 764 nt if there was a deletion of exon 52. FIG. 42 shows P2 mice injected via facial vein with AAV9 ("JA10 (P2 AAV9)") at between 36-50 hrs old, and adult mice injected via tail vein with 3.3-7.7E12 of AAV8 ("JA11 (TV AAV8)") or 4.3-7.5E12 of AAV9 ("JA12 (TV AAV9)"). As shown in FIG. 42, editing occurred to varying degrees in P2 mice (48-54 hrs old mice) AAV9 treated mice, and AAV8 and AAV9 adult treated mice, which was further confirmed by sequencing the deletion band using the primer ttctgtgattttctttggattg (SEQ ID NO: 109) which binds to exon 53. FIG. 43 shows a representative chromatogram showing the deletion of exons 51 and 52 in the sequence of the deletion band from Ja10 mouse 1.

[00244] It is understood that the foregoing detailed description and accompanying examples are merely illustrative and are not to be taken as limitations upon the scope of the invention, which is defined solely by the appended claims and their equivalents.

[00245] Various changes and modifications to the disclosed embodiments will be apparent to those skilled in the art. Such changes and modifications, including without limitation those relating to the chemical structures, substituents, derivatives, intermediates, syntheses, compositions, formulations, or methods of use of the invention, may be made without departing from the spirit and scope thereof.

[00246] For reasons of completeness, various aspects of the invention are set out in the following numbered clause:

[00247] Clause 1. A guide RNA (gRNA) comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42 or a complement thereof.

[00248] Clause 2. A DNA targeting composition comprising a first gRNA and a second gRNA, the first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9,

SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42, or a complement thereof, wherein the first gRNA molecule and the second gRNA molecule comprise different targeting domains.

[00249] Clause 3. The DNA targeting composition clause 2, wherein the first gRNA molecule is SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, or SEQ ID NO: 41, and the second gRNA molecule is SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, or SEQ ID NO: 42.

[00250] Clause 4. The DNA targeting composition of clause 2 or 3, wherein the first gRNA molecule is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14 and SEQ ID NO: 15, and the second gRNA molecule is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, and SEQ ID NO: 19.

[00251] Clause 5. The DNA targeting composition of any one of clauses 2-4, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of: (i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2; (ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (vi) a first gRNA molecule comprising a targeting domain

that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; (xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; and (xii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 41, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 42.

[00252] Clause 6. The DNA targeting composition of any one of clauses 2-5, further comprising a Clustered Regularly Interspaced Short Palindromic Repeats associated (Cas) protein.

[00253] Clause 7. The DNA targeting composition of clause 6, wherein the Cas protein comprises a Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25).

[00254] Clause 8. The DNA targeting composition of clause 6 or 7, wherein the Cas protein comprises a *Staphylococcus aureus* Cas9 molecule having an amino acid sequence of SEQ ID NO: 45.

[00255] Clause 9. The DNA targeting composition of any one of clauses 2-8, wherein the DNA targeting composition comprises a nucleotide sequence of SEQ ID NO: 83, a nucleotide

sequence of SEQ ID NO: 84, a nucleotide sequence of SEQ ID NO: 37, and/or a nucleotide sequence of SEQ ID NO: 38.

[00256] Clause 10. An isolated polynucleotide comprising the gRNA molecule of clause 1 or the DNA targeting composition of any one of clauses 2-9.

[00257] Clause 11. A vector comprising the gRNA of clause 1, the DNA targeting composition of any one of clauses 2-9, or the isolated polynucleotide of clause 10.

[00258] Clause 12. A vector comprising the DNA targeting composition of any one of clauses 6-9.

[00259] Clause 13. A vector encoding: (a) a first guide RNA (gRNA) molecule, (b) a second gRNA molecule, and (c) at least one Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), wherein the first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42, or a complement thereof, wherein the first gRNA molecule and the second gRNA molecule comprise different targeting domains.

[00260] Clause 14. The vector of clause 13, wherein the vector is configured to form a first and a second double strand break in a first and a second intron flanking exon 51 of the human DMD gene.

[00261] Clause 15. The vector of clause 13 or 14, wherein the first gRNA molecule is SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, or SEQ ID NO: 41, and the second gRNA molecule is SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, or SEQ ID NO: 42.

[00262] Clause 16. The vector of any one of clauses 13-15, wherein the first gRNA molecule is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14 and SEQ ID NO: 15, and the second gRNA molecule is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, and SEQ ID NO: 19.

[00263] Clause 17. The vector of any one of clauses 13-16, the first gRNA molecule and the second gRNA molecule are selected from the group consisting of: (i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2; (ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; and (xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second

gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18.

[00264] Clause 18. The vector of any one of clauses 11-17, wherein the vector is a viral vector.

[00265] Clause 19. The vector of clause 18, wherein the vector is an Adeno-associated virus (AAV) vector.

[00266] Clause 20. The vector of clause 19, wherein the AAV vector is an AAV8 vector or an AAV9 vector.

[00267] Clause 21. The vector of any one of clauses 11-20, wherein the vector comprises a tissue-specific promoter operably linked to the nucleotide sequence encoding the first gRNA molecule, the second gRNA molecule, and/or the Cas9 molecule.

[00268] Clause 22. The vector of clause 21, wherein the tissue-specific promoter is a muscle specific promoter.

[00269] Clause 23. A cell comprising the gRNA of clause 1, the DNA targeting composition of any one of clauses 2-9, the isolated polynucleotide of clause 10, or the vector of any one of clauses 11-22.

[00270] Clause 24. A kit comprising the gRNA of clause 1, the DNA targeting system of any one of clauses 2-9, the isolated polynucleotide of clause 10, the vector of any one of clauses 11-22, or the cell of clause 23, and optionally instructions for use.

[00271] Clause 25. A method of correcting a mutant dystrophin gene in a cell, the method comprising administering to a cell the gRNA of clause 1, the DNA targeting system of any one of clauses 2-9, the isolated polynucleotide of clause 10, or the vector of any one of clauses 11-22.

[00272] Clause 26. A method of genome editing a mutant dystrophin gene in a subject, the method comprising administering to the subject a genome editing composition comprising the gRNA of clause 1, the DNA targeting system of any one of clauses 2-9, the isolated polynucleotide of clause 10, the vector of any one of clauses 11-22, or the cell of clause 23.

[00273] Clause 27. The method of clause 26, wherein the genome editing composition is administered the subject intramuscularly, intravenously or a combination thereof.

[00274] Clause 28. The method of any one of clauses 25-27, wherein correcting the mutant dystrophin gene comprises nuclease-mediated non-homologous end joining.

[00275] Clause 29. A method of treating a subject in need thereof having a mutant dystrophin gene, the method comprising administering to the subject the gRNA of clause 1, the DNA targeting system of any one of clauses 2-9, the isolated polynucleotide of clause 10, the vector of any one of clauses 11-22, or the cell of clause 23.

[00276] Clause 30. A modified adeno-associated viral vector for genome editing a mutant dystrophin gene in a subject comprising a first polynucleotide sequence encoding the gRNA of clause 1, and a second polynucleotide sequence encoding a Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25).

[00277] Clause 31. The modified adeno-associated viral vector of clause 30, wherein the modified adeno-associated viral vector comprises the nucleotide sequence set forth in SEQ ID NO: 39 or SEQ ID NO: 40.

[00278] Clause 32. A composition for deleting a segment of a dystrophin gene comprising exon 51, the composition comprising: (a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and (b) a second vector comprising a polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), wherein each of the first and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the first vector and second vector are configured to form a first and a second double strand break in a first intron and a second intron flanking exon 51 of the human DMD gene, respectively, thereby deleting a segment of the dystrophin gene comprising exon 51.

[00279] Clause 33. The composition of clause 32, wherein the segment has a length of about 50 base pairs to about 2,000 base pairs.

[00280] Clause 34. The composition of clause 33, wherein the segment has a length of about 118 base pairs, about 233 base pairs, about 326 base pairs, about 766 base pairs, about 805 base pairs, or about 1611 base pairs.

[00281] Clause 35. The composition of any one of clauses 32-34, wherein the first Cas9 molecule and the second Cas9 molecule are the same.

[00282] Clause 36. The composition of clause 35, wherein the first Cas9 molecule and the second Cas9 molecule is a *Staphylococcus aureus* Cas9 molecule.

[00283] Clause 37. The composition of clause 36, wherein the first Cas9 molecule and the second Cas9 molecule is a mutant *Staphylococcus aureus* Cas9 molecule.

[00284] Clause 38. The composition of any one of clauses 32-34, wherein the first Cas9 molecule and the second Cas9 molecule are different.

[00285] Clause 39. The composition of clause 38, wherein the first Cas9 molecule or the second Cas9 molecule is a *Staphylococcus aureus* Cas9 molecule.

[00286] Clause 40. The composition of any one of clauses 32-39, wherein the first Cas9 molecule and/or the second Cas9 molecule comprises a SaCas9 molecule having an amino acid sequence of SEQ ID NO: 45.

[00287] Clause 41. The composition of any one of clauses 32-40, wherein the first vector and/or the second vector is a viral vector.

[00288] Clause 42. The composition of clause 41, wherein the first vector and/or the second vector is an Adeno-associated virus (AAV) vector.

[00289] Clause 43. The composition of clause 42, wherein the AAV vector is an AAV8 vector or an AAV9 vector.

[00290] Clause 44. The composition of any one of clauses 32-43, wherein the dystrophin gene is a human dystrophin gene.

[00291] Clause 45. The composition of any one of clauses 32-44, wherein the first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, or a complement thereof, wherein the first gRNA molecule and the second gRNA molecule comprise different targeting domains.

[00292] Clause 46. The composition of any one of clauses 32-45, wherein the first gRNA molecule is SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, or SEQ ID NO: 15, and the second gRNA molecule is SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, or SEQ ID NO: 19.

[00293] Clause 47. The composition of any one of clauses 32-46, wherein the first gRNA molecule is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14 and SEQ ID NO: 15, and the second gRNA molecule is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, and SEQ ID NO: 19.

[00294] Clause 48. The composition of any one of clauses 32-47, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of: (i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2; (ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (x) a first

gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; and (xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18.

[00295] Clause 49. The composition of any one of clauses 32-48, wherein the first vector comprises a nucleotide sequence set forth in SEQ ID NO: 39 and the second vector comprises a nucleotide sequence set forth in SEQ ID NO: 40.

[00296] Clause 50. The composition of any one of clauses 32-49, for use in a medicament.

[00297] Clause 51. The composition of any one of clauses 32-50, for use in the treatment of Duchenne Muscular Dystrophy.

[00298] Clause 52. A cell comprising the composition of any one of clauses 32-51.

[00299] Clause 53. A method of correcting a mutant dystrophin gene in a cell, comprising administering to the cell: (a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and (b) a second vector comprising a polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), wherein each of the first gRNA and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the vector is configured to form a first and a second double strand break in a first and a second intron flanking exon 51 of the human dystrophin gene, respectively, thereby deleting a segment of the dystrophin gene comprising exon 51.

[00300] Clause 54. The method of clause 53, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of: (i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2; (ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a

targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; and (xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18.

[00301] Clause 55. The method of clause 53 or 54, wherein the mutant dystrophin gene comprises a premature stop codon, disrupted reading frame, an aberrant splice acceptor site, or an aberrant splice donor site.

[00302] Clause 56. The method of any one of clauses 53-55, wherein the mutant dystrophin gene comprises a frameshift mutation which causes a premature stop codon and a truncated gene product.

[00303] Clause 57. The method of any one of clauses 53-55, wherein the mutant dystrophin gene comprises a deletion of one or more exons which disrupts the reading frame.

[00304] Clause 58. The method of any one of clauses 53-57, wherein the correction of the mutant dystrophin gene comprises a deletion of a premature stop codon, correction of a disrupted reading frame, or modulation of splicing by disruption of a splice acceptor site or disruption of a splice donor sequence.

[00305] Clause 59. The method of any one of clauses 53-58, wherein the correction of the mutant dystrophin gene comprises deletion of exon 51.

[00306] Clause 60. The method of any one of clauses 53-59, wherein the correction of the mutant dystrophin gene comprises homology-directed repair.

[00307] Clause 61. The method of clause 60, further comprising administering to the cell a donor DNA.

[00308] Clause 62. The method of any one of clauses 53-61, wherein the correction of the mutant dystrophin gene comprises nuclease mediated non-homologous end joining.

[00309] Clause 63. The method of any one of clauses 53-62, wherein the cell is a myoblast cell.

[00310] Clause 64. The method of any one of clauses 53-63, wherein the cell is from a subject suffering from Duchenne muscular dystrophy.

[00311] Clause 65. The method of any one of clauses 53-64, wherein the cell is a myoblast from a human subject suffering from Duchenne muscular dystrophy.

[00312] Clause 66. The method of any one of clauses 53-65, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of: (i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2; (ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; and (iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19.

[00313] Clause 67. A method of treating a subject in need thereof having a mutant dystrophin gene, the method comprising administering to the subject: (a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and (b) a second vector comprising a polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), wherein each of the first gRNA and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the vector is configured to form a first and a second double strand break in a first and a second intron flanking exon 51 of the human dystrophin gene, respectively, thereby deleting a segment of the dystrophin gene comprising exon 51.

[00314] Clause 68. The method of clause 67, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of: (i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2; (ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second

gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; (viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; (ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19; (x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; and (xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18.

[00315] Clause 69. The method of clause 68, wherein the subject is suffering from Duchenne muscular dystrophy.

[00316] Clause 70. The method of any one of clauses 67-69, wherein the first vector and second vector are administered to a muscle of the subject.

[00317] Clause 71. The method of clause 70, wherein the muscle is skeletal muscle or cardiac muscle.

[00318] Clause 72. The method of clause 71, wherein the skeletal muscle is tibialis anterior muscle.

[00319] Clause 73. The method of any one of clauses 67-72, wherein the first vector and second vector are administered to the subject intramuscularly, intravenously or a combination thereof.

[00320] Clause 74. A method of generating a transgenic rodent embryo having a human dystrophin gene (hDMD) with an exon 52 deletion ($\Delta 52$), the method comprising administering to a rodent embryo the gRNA of clause 1, the DNA targeting system of any one of clauses 2-9, the isolated polynucleotide of clause 10, the vector of any one of clauses 11-22, the modified adeno-associated viral vector of clause 30 or 31, or the composition of any one of clauses 32-51, thereby deleting exon 52 of the human dystrophin gene, and selecting for a transgenic rodent embryo having a deletion of exon 52 of the human dystrophin gene, wherein the rodent embryo comprises a normal human dystrophin gene.

[00321] Clause 75. The method of clause 74, wherein the rodent embryo is a mouse embryo.

[00322] Clause 76. The method of clause 74 or 75, wherein the transgenic rodent embryo is heterozygous hDMD or heterozygous hDMD-Δ52.

[00323] Clause 77. The method of any one of clauses 74-76, wherein a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 41, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 42 are administered to the rodent embryo to delete exon 52 of the human dystrophin gene.

[00324] Clause 78. The method of any one of clauses 74-77, further comprising administering to the rodent embryo a Cas protein comprising an amino acid sequence set forth in SEQ ID NO: 27.

[00325] Clause 79. A transgenic rodent embryo produced by the method of any one of clauses 74-78.

[00326] Clause 80. A transgenic rodent produced from the transgenic rodent embryo of clause 79.

Appendix

pDO240 with JCR179 (SEQ ID NO: 37) (gRNA in bold)

ctaaattgtaagcgtaatatattttgttaaaattcgcgttaaattttgttaaatcagctcatttttaaccaatagg
 ccgaaatcggcaaaatcccttataaatcaaaagaatagaccgagatagggttgagtgttccagtttggacaag
 agtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatcaggcgatggcccactacgtga
 accatcacctaatacaagtttttggggcgcaggtgccgtaaaagcactaaatcggaaccctaaaggagccccgat
 ttagagcttgacggggaaagccggcgaaacgtggcgagaaagggaaggaaagcgaaaaggagcgggcgctaggggcg
 ctggcaagtgtagcggcgcgtgcgtaaccaccacacccgccgcgttaatgcgccgtacagggcgcgctccca
 ttgccattcaggctgcgcaactgttgggaagggcgatcggcgccgtcttcgctattacgccagctggcgaaag
 ggggatgtgctgcaaggcgattaagttgggtaacgccaggggtttccagtcacgacgttgtaaacgacggccagt
 gagcgcgcgtaatacgaactactataggcgcaattgggtaccaagcttgcctatttccatgattccttcattttg
 catatacgatacaaggctgttagagagataattggaattaattgactgtaaacacaaagatattagtacaaaatac
 gtgacgtagaaaagtaataatttctgggtagtttgagttttaaattatgttttaaatggactatcatatgctta
 ccgtaactgaaagtatttcgatttctggctttatatcttggtaaggacgaaacacc**AACACACAGCTGGGT**
TATCAGAGgttttagtactctggaaacagaatctactaaaacaaggcaaatgccgtgtttatctcgtcaactgtt
 ggcgagattttttgcggccgcccgcggtggagctccagcttttgtcccttagtgagggtaattgcgcgcttgg
 cgtaatcatggcatagctgttctgtgtgaaattgtatccgctcacaattccacacaacatacgagccggaagc
 ataaagtgtaaagcctggggcgcctaatagtgagctaaactcacattaattgcgttgcgctcactgccgcttcca
 gtccgggaaacctgtcgtgccagctgcattaatgaatcgccaacgcgcggggagaggcggtttgcgtattgggcgct
 ctccgcttctcgtcactgactcgtcgcctcggtcgttcggctgcggcgagcgggtatcagctcactcaaaggcg
 gtaatacggttatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaa
 ccgtaaaaaggccgctgtgctggcgttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaa
 gtcagaggtggcgaaacccgacaggactataagataaccaggcgtttccccctggaagctccctcgtgcgctctct
 gttccgacctgccgcttaccggatacctgtccgcttttcccttcgggaagcgtggcgctttctcatagctcacg
 ctgtaggatctcagttcgggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccgacc
 gctgcgccttatccgtaactatcgtcttgagtccaacccggttaagacacgacttatcgccactggcagcagccact
 ggtaacaggattagcagagcgaggtatgtagggcgtgtacagagttcttgaagtggcctaactacggctacac
 tagaaggacagtatttggatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccg
 gcaaacaaaccaccgctggtagcgggtgggtttttgttgcaagcagcagattacgcgcagaaaaaaggatctcaa
 gaagatcctttgatctttctacggggctgacgctcagtggaacgaaaactcacgttaagggttttggatcatgag
 attatcaaaaaggatcttcacctagatccttttaattaaaaatgaagttttaaataaatctaaagtatatatgagt

aaacttggtctgacagttaccaatgcttaatcagtgaggcacctatctcagcgatctgtctatttcgttcacata
gttgccctgactccccgtcgtgtgataactacgatacgggagggccttaccatctggccccagtgctgcaatgatacc
gcgagacccacgctcaccgggtccagatttatcagcaataaaccagccagccggaagggccgagcgcagaagtggct
ctgcaactttatccgcctccatccagcttattaattgttgccgggaagctagagtaagtagttcgccagttaatagt
ttgcgcaacggttggccattgctacaggcatcggtgtgtcacgctcgtcgtttgggtatggcttcattcagctccgg
ttccaacgatcaaggcgagttacatgatccccatgtgtgcaaaaaagcggttagctccttcggctcctccgatcg
ttgtcagaagtaagttggccgagtggtatcactcatggttatggcagcactgcataattcttactgtcatgcc
tccgtaagatgcttttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatcgggcgaccgagttg
ctcttggccggcgtaatacgggataataccgcgccacatagcagaactttaaaagtgtcatcattggaaaacgtt
cttcggggcgaaaaactcaaggatcttaccgctgttgagatccagttcgatgtaaccactcgtgcacccaactga
tcttcagcatcttttactttcaccagcggttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaagggaat
aagggcgacacggaaatgtgaatactcatactcttccttttcaatattattgaagcatttatcagggtattgtc
tcatgagcggatacatatttgaatgtatttagaaaaataacaaataggggttccgcgcacatttccccgaaaagtg ccac

pDO240 with JCR183 (SEQ ID NO: 38) (gRNA in bold)

ctaaattgtaagcgtaataattttgttaaaattcgcgttaaattttgttaaatcagctcatttttaaccaatagg
ccgaaatcggaataatccctataaatcaaaagaatagaccgagataggggtgagtggttccagtttggaacaag
agtccactattaaagaacgtggactccaacgtcaaggggcgaaaaaccgtctatcagggcgatggccactacgtga
accatcacctaataagtttttggggtcgaggtgccgtaaaagcactaaatcggaaccctaaaggagccccgat
ttagagcttgacggggaaagccggcgaaacgtggcgagaaagggaagggaagaaagcgaaaggagcggcgctagggcg
ctggcaagtgtagcgggtcacgctgcgcgtaaccaccacacccgccgcgttaatgcgccgtacagggcgcgctccca
ttcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgtattacgccagctggcgaaag
ggggatgtgtgcaaggcgattaagttgggtaacgccagggttttccagtcacgacgttgtaaaacgacggccagt
gagcgcgcgtaatacgaactcactatagggcgaaattgggtaccaagcttgccatttccatgattccttcataattg
catatacgatacaaggctgttagagagataattggaattaatttgactgtaaacacaaagatattagtacaaaatac
gtgacgtagaaagtaataatttctgggtagttgcagttttaaattatgttttaaattggactatcatatgctta
ccgtaacttgaaagtatttcgatttcttggtttatatacttggtaaaggacgaaacacc**AGAACTGGTGGGAAA**
TGGTCTAGgttttagtactctggaacagaatctactaaaacaaggcaaaatgccgtgtttatctcgtcaactgtt
ggcgagattttttgcggccgcccggtggagctccagctttgttcccttagtgagggtaattgcgcgcttg
cgtaatcatggcatagctgttctgtgtgaaattgtatccgctcacaattccacacacatacgagccggaagc
ataaagtgtaaagcctgggggtgcctaatagtgagtaactcacattaattgcgttcgctcactgccgctttcca

gtcgggaaacctgtcgtgccagctgcattaatgaatggccaacgcgcggggagaggcggtttgcgtattgggcgct
 ctccgcttcctcgtcactgactcgtcgcctcggtcgttcgggtgcggcgagcgggtatcagctcactcaaaggcg
 gtaatacgggtatccacagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaa
 ccgtaaaaaggccggttgctggcgtttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaa
 gtcagagggtggcgaaacccgacaggactataaagataccaggcggtttccccctggaagctccctcgtcgcgtctcct
 gttccgacctgcccgttaccggatacctgtccgctttctcccttcgggaagcgtggcgctttctcatagctcacg
 ctgtaggtatctcagttcgggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagcccagacc
 gctgcgccttatccgtaactatcgtcttgagtccaacccgtaagacacgacttatgccactggcagcagccact
 ggtaacaggattagcagagcgagggtatgtaggcggtgctacagagttctgaagtgggtggcctaactacggctacac
 tagaaggacagtatttggtatctgcgctctgctgaagccagttaccttcggaaaaagagttggtagctcttgatccg
 gcaacaaaccaccgctggtagcgggtggtttttgttgcaagcagcagattacgcgcagaaaaaaggatctcaa
 gaagatcctttgatctttctacggggtctgacgctcagtggaacgaaaactcacgttaagggttttggtcatgag
 attatcaaaaaggatcttcacctagatccttttaataaaaaatgaagttttaaatcaatctaaagtatatatgagt
 aaacttggtctgacagttaccaatgcttaacagtgaggcacctatctcagcgatctgtctatttcgtcatccata
 gttgcctgactccccgtcgtgtagataactacgatacgggaggggttaccatctggccccagtgctgcaatgatacc
 gcgagaccacgctcaccgggtccagatttatcagcaataaaccagccagccggaaggccgagcgcagaagtggtc
 ctgcaactttatccgctccatccagttatattaattgttgcgggaagctagagtaagtagttcggcagttaatagt
 ttgcgcaacgtttgtccattgctacaggcatcgtggtgtcacgctcgtcgtttgggtatggcttcattcagctccgg
 ttccaacgatcaaggcgagttacatgatccccatgttggtgcaaaaaagcggtagctccttcggctcctccgatcg
 ttgtcagaagtaagttggccgagtggtatcactcatggttatggcagcactgcataattctctactgtcatgccaa
 tccgtaagatgctttctgtgactggtgagtactcaaccaagtcattctgagaatagtgtatgcggcgaccgagttg
 ctcttgcggcggtcaataacgggataataccgcgccacatagcagaactttaaaagtgtcatcattggaaaacgtt
 cttcggggcgaaaactctcaaggatcttaccgctgttgagatccagttcagatgaacccactcgtgcacccaactga
 tcttcagcatctttactttaccagcgtttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggaat
 aaggcgacacggaaatgttgaatactcatactcttcttttcaatattattgaagcatttatcagggttatgtc
 tcatgagcggatacatattgaatgtatttagaaaaataaacaatagggttccgcgcacattccccgaaaagtgcac

PT366 with JCR179 PT366AAV 179 (SEQ ID NO: 39) – AAV plasmid used for in vivo work (gRNA in bold; SaCas9 is uppercase; NLS is lowercase, bold, and underlined)

cctgcaggcagctgcgcgctcgtcgtcactgaggccgccccgggcgtcgggcgacctttggtcggcgccctcagt
 gagcgagcgagcgcgcagagagggtggccaactccatcactagggttctcggcctctagactcagggcgttg

acattgattattgactagtattaatagtaatacaattacgggggtcattagttcatagcccatatatggagttccgcg
ttacataacttacggtaaatggcccgctggctgaccgccaacgacccccgccattgacgtcaataatgacgtat
gttcccatagtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaacgccacttggc
agtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggcccgctggcattatg
cccagtacatgaccttatgggactttcctacttggcagtacatctacgtattagtcacgtattaccatggtgatg
cggttttggcagttacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccattgacgt
caatgggagttgttttggcaccaaatcaacgggactttccaaatgtcgtaacaactccgccccattgacgcaa
tgggcggtaggcgtgtacgggtgggaggtctatataagcagagctctctggctaactaccggtgccaccatggcccca
aagaagaagcgggaaggctcggtatccacggagtcacagcagccAAGCGGA ACTACATCCTGGGCCTGGACAT
CGGCATCACCAGCGTGGGCTACGGCATCATCGACTACGAGACACGGGACGTGATCG
ATGCCGGCGTGCGGCTGTTCAAAGAGGCCAACGTGGAAAACAACGAGGGCAGGCG
GAGCAAGAGAGGCGCCAGAAGGCTGAAGCGGCGGAGGCGGCATAGAATCCAGAGA
GTGAAGAAGCTGCTGTTTCGACTACAACCTGCTGACCGACCACAGCGAGCTGAGCGG
CATCAACCCCTACGAGGCCAGAGTGAAGGGCCTGAGCCAGAAGCTGAGCGAGGAA
GAGTTCTCTGCCGCCCTGCTGCACCTGGCCAAGAGAAAGGCGTGCACAACGTGAAC
GAGGTGGAAGAGGACACCGGCAACGAGCTGTCCACCAAAGAGCAGATCAGCCGGA
ACAGCAAGGCCCTGGAAGAGAAATACGTGGCCGA ACTGCAGCTGGAACGGCTGAA
GAAAGACGGCGAAGTGCGGGGCAGCATCAACAGATTCAAGACCAGCGACTACGTG
AAAGAAGCCAAACAGCTGCTGAAGGTGCAGAAGGCCTACCACCAGCTGGACCAGA
GCTTCATCGACACCTACATCGACCTGCTGGAAACCCGGCGGACCTACTATGAGGGA
CCTGGCGAGGGCAGCCCCTTCGGCTGGAAGGACATCAAAGAATGGTACGAGATGCT
GATGGGCCACTGCACCTACTTCCCCGAGGAACTGCGGAGCGTGAAGTACGCCTACA
ACGCCGACCTGTACAACGCCCTGAACGACCTGAACAATCTCGTGATCACCAGGGAC
GAGAACGAGAAGCTGGAATATTACGAGAAGTTCCAGATCATCGAGAACGTGTTCAA
GCAGAAGAAGAAGCCACCCTGAAGCAGATCGCCAAAGAAATCCTCGTGAACGAA
GAGGATATTAAGGGCTACAGAGTGACCAGCACCGGCAAGCCGAGTTCACCAACCTG
AAGGTGTACCACGACATCAAGGACATTACCGCCCGGAAAGAGATTATTGAGAACGC
CGAGCTGCTGGATCAGATTGCCAAGATCCTGACCATCTACCAGAGCAGCGAGGACA
TCCAGGAAGAACTGACCAATCTGAACTCCGAGCTGACCCAGGAAGAGATCGAGCAG
ATCTCTAATCTGAAGGGCTATAACGGCACCCACAACCTGAGCCTGAAGGCCATCAA
CCTGATCCTGGACGAGCTGTGGCACACCAACGACAACCAGATCGCTATCTTCAACCG

GCTGAAGCTGGTGCCCAAGAAGGTGGACCTGTCCCAGCAGAAAGAGATCCCCACCA
CCCTGGTGGACGACTTCATCCTGAGCCCCGTCGTGAAGAGAAGCTTCATCCAGAGCA
TCAAAGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAACGACATCATTATC
GAGCTGGCCCCGCGAGAAGAAGTCCAAGGACGCCCAGAAAATGATCAACGAGATGC
AGAAGCGGAACCGGCAGACCAACGAGCGGATCGAGGAAATCATCCGGACCACCGG
CAAAGAGAACGCCAAGTACCTGATCGAGAAGATCAAGCTGCACGACATGCAGGAA
GGCAAGTGCCTGTACAGCCTGGAAGCCATCCCTCTGGAAGATCTGCTGAACAACCC
CTTCAACTATGAGGTGGACCACATCATCCCCAGAAGCGTGTCTTCGACAACAGCTT
CAACAACAAGGTGCTCGTGAAGCAGGAAGAAAACAGCAAGAAGGGCAACCGGACC
CCATTCCAGTACCTGAGCAGCAGCGACAGCAAGATCAGCTACGAAACCTTCAAGAA
GCACATCCTGAATCTGGCCAAGGGCAAGGGCAGAATCAGCAAGACCAAGAAAGAG
TATCTGCTGGAAGAACGGGACATCAACAGGTTCTCCGTGCAGAAAGACTTCATCAA
CCGGAACCTGGTGGATACCAGATACGCCACCAGAGGCCTGATGAACCTGCTGCGGA
GCTACTTCAGAGTGAACAACCTGGACGTGAAAGTGAAGTCCATCAATGGCGGCTTC
ACCAGCTTTCTGCGGCGGAAGTGGAAGTTTAAGAAAGAGCGGAACAAGGGGTACAA
GCACCACGCCGAGGACGCCCTGATCATTGCCAACGCCGATTTTCATCTTCAAAGAGTG
GAAGAACTGGACAAGGCCAAAAAAGTGATGGAAAACCAGATGTTCGAGGAAAAG
CAGGCCGAGAGCATGCCCGAGATCGAAACCGAGCAGGAGTACAAAGAGATCTTCAT
CACCCCCACCAGATCAAGCACATTAAGGACTTCAAGGACTACAAGTACAGCCACC
GGGTGGACAAGAAGCCTAATAGAGAGCTGATTAACGACACCCTGTACTCCACCCGG
AAGGACGACAAGGGCAACACCCTGATCGTGAACAATCTGAACGGCCTGTACGACAA
GGACAATGACAAGCTGAAAAAGCTGATCAACAAGAGCCCCGAAAAGCTGCTGATGT
ACCACCACGACCCCCAGACCTACCAGAACTGAAGCTGATTATGGAACAGTACGGC
GACGAGAAGAATCCCCTGTACAAGTACTACGAGGAAACCGGGAACCTACCTGACCAA
GTACTCCAAAAAGGACAACGGCCCCGTGATCAAGAAGATTAAGTATTACGGCAACA
AACTGAACGCCCATCTGGACATCACCGACGACTACCCCAACAGCAGAAACAAGGTC
GTGAAGCTGTCCCTGAAGCCCTACAGATTCGACGTGTACCTGGACAATGGCGTGTAC
AAGTTCGTGACCGTGAAGAATCTGGATGTGATCAAAAAAGAAAAGTACTACGAAGT
GAATAGCAAGTGCTATGAGGAAGCTAAGAAGCTGAAGAAGATCAGCAACCAGGCC
GAGTTTATCGCCTCCTTCTACAACAACGATCTGATCAAGATCAACGGCGAGCTGTAT
AGAGTGATCGGCGTGAACAACGACCTGCTGAACCGGATCGAAGTGAACATGATCGA

CATCACCTACCGCGAGTACCTGGAAAACATGAACGACAAGAGGCCCCCCCAGGATCA
TTAAGACAATCGCCTCCAAGACCCAGAGCATTAAAGAAGTACAGCACAGACATTCTG
GGCAACCTGTATGAAGTGAAATCTAAGAAGCACCCCTCAGATCATCAAAAAGGGC~~aaa~~

aggccggcgccacgaaaaaggccggccaggca

aaaaagaaaaagggcatcctaccatacgaatgtccagattacgcttaccatacgaatgtccagattacgcttacc

atagatgtccagattacgcttaagaattcctagagctcgtgatcagcctcgactgtgccttctagtgtccagcc

atctgtgtttgccctccccctgaccttctgacctggaaggtgccactccactgtcctttcctaataaaatg

aggaaattgcatcgattgtctgagtaggtgtcattctattctggggggtggggtggggcaggacagcaagggggag

gattgggaagagaatagcaggcatgctggggagggtaccgagggcctatttcccatgattcctcatatttgcata

cgatacaaggctgttagagagataattggaattaatttgactgtaaacacaaagatattagtacaaaatacgtgacg

tagaaagtaataatttctgggtagttgcagtttaaaattatgtttaaatggactatcatatgcttaccgtaa

cttgaaagtatttcgatttcttggtttatatacttctgtgaaaggacgaaacaccg**AACACACAGCTGGGTTATCA**

GAGgttttagtactctggaacagaatctactaaaacaaggcaaaatgccgtgtttatctcgtcaactgttggcga

gatttttgcggccgcaggaaccctagtgtgaggtggccactccctctctgcgcgctcgtcgtcactgagggc

gggcgaccaaaaggctgcccgcagcccgggctttgcccggggcgccctcagtgcgcgagcgcgcagctgcctgca

ggggcgctgatgcggtattttctccttacgcatctgtgcggtatttcacaccgcatacgtcaaagcaaccatagta

cgcgccctgtagcggcgcatgaagcgcggcggtgtggtggttacgcgcagcgtgaccgctacacttgccagcgccc

tagcggccgctcctttcgtttcttcccttcttctcgccacgttcgccggctttcccgtcaagctctaaatcgg

gggctcccttagggttccgatttagtgctttacggcacctcgaccccaaaaacttgatttgggtgatggttcacg

tagtgggcatcgccctgatagacgggttttgcctttgacgttggagtccacgttcttaatagtggactcttgt

tccaaactggaacaacactcaaccctatctcgggctattctttgattataagggattttgccgatttcggcctat

tgggtaaaaaatgagctgatttaacaaaaatttaacgcgaatttaacaaaatattaacgtttacaattttatggtg

cactctcagtacaatctgctctgatgccgatagttaagccagccccgacacccgccaacacccgctgacgcgcct

gacgggcttgtctgctccggcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtt

tcaccgtcatcaccgaaacgcgcgagacgaaaggcctcgtgatacgcctattttataggtaatgtcatgataat

aatggtttcttagacgtcaggtggcacttttcggggaaatgtgcgcggaacccctatttgttttttctaaatac

attcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataatattgaaaaaggagagatgagat

attcaacatttcggtgctgcccttattccctttttgcggcattttgccttctgttttctcaccagaaacgct

ggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggttacatgaactggatctcaacagcggtaga

tccttgagagttttgccccgaagaacgtttccaatgatgagcacttttaagttctgctatgtggcgcggtatta

tcccgattgacgcggggaagagcaactcggctgccgcatacactattctcagaatgacttggtgagtactacc

agtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgtgccataacatgagtataaca
 ctgcgcccaacttacttctgacaacgatcggaggaccgaaggagctaaccgctttttgcacaacatgggggatcat
 gtaactcgccttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgt
 agcaatggcaacaacgttgcgcaaaactattaactggcgaactacttactctagcttcccgcaacaattaatagact
 ggatggagggcgataaagtgcaggaccacttctgcgctcggcccttcggctggctggtttattgctgataaatct
 ggagccggtgagcgtggaagccgcggtatcattgcagcactggggccagatggttaagccctcccgatcgtagtatt
 ctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgtgagatagggtgcctcactgattaagc
 attgtaactgtcagaccaagtttactcatatatactttagattgatttaaaacttcatttttaatttaaaggatc
 taggtgaagatccttttgataatctcatgacaaaaatccctaactgagtttctgtccactgagcgtcagaccc
 cgtagaaaagatcaaaggatcttctgagatcctttttctgcgcgtaatctgctgcttgcacaacaaaaaacac
 cgctaccagcgggtggtttgttgcggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcg
 cagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacata
 cctcgctctgtaatcctgttaccagtggctgctgccagtggcgataagtcgtgtcttaccgggttgactcaagac
 gatagttaccggataaggcgacgggtcgggtgaacggggggttcgtgcacacagcccagcttgagcgaacgacc
 tacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgcttccgaaggagaaaggcggacaggta
 tccggttaagcggcagggtcggaacaggagagcgcacgaggagcttccaggggaaacgcctggtatctttatagtc
 ctgtcgggtttcgccacctctgacttgagcgtcgattttgtgatgctcgtcagggggcgagcctatggaaaaac
 gccagcaacgcggccttttacggttcttgcccttttgctggcctttgctcacatgt

PT366 with JCR183 (SEQ ID NO: 40) – used for in vivo work (gRNA in bold; SaCas9 is uppercase; NLS is lowercase, bold, and underlined)

cctgcaggcagctgcgcgctcgtcgtcactgaggccgcccggcgctggggcgaccttggcgcccgccctcagt
 gagcgagcgagcgcgagagaggagtgccaactccatcactagggggtcctgcggcctctagactcgaggcggtg
 acattgattattgactagtattataatagtaatcaattacgggggtcattagttcatagcccatatatggagtccgcg
 ttacataacttacggtaaatggccgcctggctgaccgccaacgacccccgccattgacgtcaataatgacgtat
 gttcccatagtaacgccaatagggactttcattgacgtcaatgggtggagtatttacggtaactgcccacttggc
 agtacatcaagtgtatcatatgccaagtacgccccctattgacgtcaatgacggtaaatggccgcctggcattatg
 ccagttacatgacctatgggactttctacttggcagttacatctacgtattagtcacgtattaccatggtgatg
 cggttttggcagttacatcaatgggcgtggatagcgggttgactcacggggattccaagtctccacccattgacgt
 caatgggagttgttttggcaccaaaatcaacgggactttccaaaatgtcgttaacaactccgccccattgacgcaa
 tggggcgtaggcgtgtacgggtgggaggtctatataagcagagctctctggctaactaccggtgccaccatggcccca

aagaagaagcggaaggtcggtatccacggaggtcccagcagccAAGCGGAAC TACATCCTGGGCCTGGACAT
CGGCATCACCAGCGTGGGCTACGGCATCATCGACTACGAGACACGGGACGTGATCG
ATGCCGGCGTGCGGCTGTTCAAAGAGGCCAACGTGGAAAACAACGAGGGCAGGCG
GAGCAAGAGAGGGCGCCAGAAGGCTGAAGCGGCGGAGGCGGCATAGAATCCAGAGA
GTGAAGAAGCTGCTGTTTCGACTACAACCTGCTGACCGACCACAGCGAGCTGAGCGG
CATCAACCCCTACGAGGCCAGAGTGAAGGGCCTGAGCCAGAAGCTGAGCGAGGAA
GAGTTCTCTGCCGCCCTGCTGCACCTGGCCAAGAGAAGAGGGCGTGCACAACGTGAA
CGAGGTGGAAGAGGACACCGGCAACGAGCTGTCCACCAAAGAGCAGATCAGCCGG
AACAGCAAGGCCCTGGAAGAGAAATACGTGGCCGAAC TGCAGCTGGAACGGCTGA
AGAAAGACGGCGAAGTGCGGGGCAGCATCAACAGATTCAAGACCAGCGACTACGT
GAAAGAAGCCAAACAGCTGCTGAAGGTGCAGAAGGCCTACCACCAGCTGGACCAG
AGCTTCATCGACACCTACATCGACCTGCTGGAAACCCGGCGGACCTACTATGAGGG
ACCTGGCGAGGGCAGCCCCCTTCGGCTGGAAGGACATCAAAGAATGGTACGAGATGC
TGATGGGCCACTGCACCTACTTCCCCGAGGAACTGCGGAGCGTGAAGTACGCCTAC
AACGCCGACCTGTACAACGCCCTGAACGACCTGAACAATCTCGTGATCACCAGGGA
CGAGAACGAGAAGCTGGAATATTACGAGAAGTTCCAGATCATCGAGAACGTGTTCA
AGCAGAAGAAGAAGCCCACCCTGAAGCAGATCGCCAAAGAAATCCTCGTGAACGA
AGAGGATATTAAGGGCTACAGAGTGACCAGCACCGGCAAGCCCGAGTTCACCAACC
TGAAGGTGTACCACGACATCAAGGACATTACCGCCCGGAAAGAGATTATTGAGAAC
GCCGAGCTGCTGGATCAGATTGCCAAGATCCTGACCATCTACCAGAGCAGCGAGGA
CATCCAGGAAGAACTGACCAATCTGAACTCCGAGCTGACCCAGGAAGAGATCGAGC
AGATCTCTAATCTGAAGGGCTATACCGGCACCCACAACCTGAGCCTGAAGGCCATC
AACCTGATCCTGGACGAGCTGTGGCACACCAACGACAACCAGATCGCTATCTTCAA
CCGGCTGAAGCTGGTGCCCAAGAAGGTGGACCTGTCCCAGCAGAAAGAGATCCCCA
CCACCCTGGTGGACGACTTCATCCTGAGCCCCGTCGTGAAGAGAAGCTTCATCCAGA
GCATCAAAGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAACGACATCATT
ATCGAGCTGGCCCGCGAGAAGAACTCCAAGGACGCCCAGAAAATGATCAACGAGAT
GCAGAAGCGGAACCGGCAGACCAACGAGCGGATCGAGGAAATCATCCGGACCACC
GGCAAAGAGAACGCCAAGTACCTGATCGAGAAGATCAAGCTGCACGACATGCAGG
AAGGCAAGTGCCTGTACAGCCTGGAAGCCATCCCTCTGGAAGATCTGCTGAACAAC
CCCTTCAACTATGAGGTGGACCACATCATCCCCAGAAGCGTGTCTTCGACAACAGC

TTCAACAACAAGGTGCTCGTGAAGCAGGAAGAAAACAGCAAGAAGGGCAACCGGA
 CCCCATTCCAGTACCTGAGCAGCAGCGACAGCAAGATCAGCTACGAAACCTTCAAG
 AAGCACATCCTGAATCTGGCCAAGGGCAAGGGCAGAATCAGCAAGACCAAGAAAG
 AGTATCTGCTGGAAGAACGGGACATCAACAGGTTCTCCGTGCAGAAAGACTTCATC
 AACCGGAACCTGGTGGATACCAGATACGCCACCAGAGGCCTGATGAACCTGCTGCG
 GAGCTACTTCAGAGTGAACAACCTGGACGTGAAAGTGAAGTCCATCAATGGCGGCT
 TCACCAGCTTTCTGCGGCGGAAGTGGAAGTTTAAGAAAGAGCGGAACAAGGGGTAC
 AAGCACCACGCCGAGGACGCCCTGATCATTGCCAACGCCGATTTTCATCTTCAAAGA
 GTGGAAGAACTGGACAAGGCCAAAAAAGTGATGGAAAACCAGATGTTCGAGGAA
 AAGCAGGCCGAGAGCATGCCCAGATCGAAACCGAGCAGGAGTACAAAGAGATCT
 TCATCACCCCCCACCAGATCAAGCACATTAAGGACTTCAAGGACTACAAGTACAGC
 CACCGGGTGGACAAGAAGCCTAATAGAGAGCTGATTAACGACACCCTGTACTCCAC
 CCGGAAGGACGACAAGGGCAACACCCTGATCGTGAACAATCTGAACGGCCTGTACG
 ACAAGGACAATGACAAGCTGAAAAAGCTGATCAACAAGAGCCCCGAAAAGCTGCT
 GATGTACCACCACGACCCCCAGACCTACCAGAACTGAAGCTGATTATGGAACAGT
 ACGGCGACGAGAAGAATCCCCTGTACAAGTACTACGAGGAAACCGGGAACCTACCTG
 ACCAAGTACTCCAAAAAGGACAACGGCCCCGTGATCAAGAAGATTAAGTATTACGG
 CAACAACTGAACGCCCATCTGGACATCACCGACGACTACCCCAACAGCAGAAACA
 AGGTCGTGAAGCTGTCCCTGAAGCCCTACAGATTCGACGTGTACCTGGACAATGGCG
 TGTACAAGTTCGTGACCGTGAAGAATCTGGATGTGATCAAAAAAGAAAACCTACTAC
 GAAGTGAATAGCAAGTGCTATGAGGAAGCTAAGAAGCTGAAGAAGATCAGCAACC
 AGGCCGAGTTTATCGCCTCCTTCTACAACAACGATCTGATCAAGATCAACGGCGAGC
 TGTATAGAGTGATCGGCGTGAACAACGACCTGCTGAACCGGATCGAAGTGAACATG
 ATCGACATCACCTACCGCGAGTACCTGGAAAACATGAACGACAAGAGGCCCCCAG
 GATCATTAAAGACAATCGCCTCCAAGACCCAGAGCATTAAAGAAGTACAGCACAGACA
 TTCTGGGCAACCTGTATGAAGTGAAATCTAAGAAGCACCTCAGATCATCAAAAAG
 GGCaaaaggccggcgccacgaaaaaggccggccaggca

aaaaagaaaaagggcatcctaccatac gatgttcagattacgcttaccatac gatgttcagattacgcttacc

atagatgttcagattacgcttaagaattcctagagctcgctgatcagcctcgactgtgccttctagtgcagcc

atctgtgtgttccccccccgtgccttcttgacctggaaggtgccactccccactgtccttcttaataaaatg

aggaaattgcatcgattgtctgagtaggtgtcattctattctgggggggtgggggtggggcaggacagcaagggggag

gattgggaagagaatagcaggcatgctggggaggtaccgagggcctatttcccatgattccttcatatttgcata
cgatacaaggctgtagagagataattggaattaattgactgtaaacacaaagatattagtacaaaatcgtgacg
tagaaagtaataatttcttgggtagtttgcagttttaaattatgttttaaattggactatcatatgcttaccgtaa
cttgaaagtatttcgatttcttggctttatatatcttgggaaaggacgaaacaccg**AGAACTGGTGGGAAATGGTC**
TAGgttttagtactctggaaacagaatctactaaaacaaggcaaatgccgtgttatctcgtcaacttgttggcga
gatttttgcggccgcaggaacccctagtgatggagttggccactccctctctgcgcgctcgtcgtcactgagggc
gggcgacaaaaggctgcccgcacccgggctttggcggggggcctcagtgagcgagcgagcgcgagctgcctgca
ggggcgctgatgcggtattttctcctacgcacatctgtgcggtatttcacaccgcatacgtcaaagcaaccatagta
cgcgccctgtagcggcgcatgaagcgcggggggtgtgtgtgttacgcgcagcgtgaccgctacacttgcgcgccc
tagcggccgctcctttcgtttcttcccttcttctcgcacggtcggcggtttcccgctcaagctctaaatcgg
gggctcccttaggggtccgatttagtgctttacggcacctcgaccccaaaaacttgatttgggtgatggttcacg
tagtgggccatcgccctgatagacgggttttcgccccttgacgttggagtccacgttcttaatagtggactcttgt
tccaaactggaacaacactcaaccctatctcgggctattctttgattataagggttttgcgatttcggcctat
tggttaaaaaatgagctgatttaacaaaatttaacgcgaatttaacaaaatattaacgtttacaattttatggtg
cactctcagtacaatctgctctgatgccgcatagttaagccagccccgacacccgccaacacccgctgacgcgcct
gacgggcttctgctcccgcatccgcttacagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtt
tcaccgtcatcaccgaaacgcgcgagacgaaagggcctcgtgatacgcctattttataggtaatgtcatgataat
aatggtttcttagacgtcaggtggcacttttcggggaatgtgcgcggaacccctatttgtttattttctaaatac
attcaaatatgtatccgctcatgagacaataaccctgataaatgcttcaataattgaaaaaggaagagtatgagt
attcaacatttccgtgtcgccttattccctttttgcggcattttgccttctgttttctcaccagaaacgct
gggtgaaagtaaaagatgctgaagatcagttgggtgcacgagtggtttacatcgaaactggatctcaacagcgtaaga
tccttgagagttttgccccgaagaacgttttcaatgatgagcacttttaagtctgctatgtggcgcggtatta
tcccgattgacgcccgggaagagcaactcggctgccgcatacactattctcagaatgacttgggtgagtactacc
agtcacagaaaagcatcttacggatggcatgacagtaagagaattatgcagtgtgccataacctgagtataaca
ctgcccgaacttacttctgacaacgatcggaggaccgaaggagctaaccgctttttgcacaacatgggggatcat
gtaactcgcttgatcgttgggaaccggagctgaatgaagccataccaaacgacgagcgtgacaccacgatgcctgt
agcaatggcaacaacgttgcgcaactattaactggcgaaactacttactctagcttccggcaacaattaatagact
ggatggaggcggaataaagtgcaggaccacttctgcgtcggccctccggctggctggtttattgctgataaatct
ggagccgggtgagcgtggaagccggtatcattgcagcactggggccagatggttaagccctcccgatcgtagttat
ctacacgacggggagtcaggcaactatggatgaacgaaatagacagatcgtgagataggtgcctcactgattaagc
attgtaactgtcagaccaagtttactatatactttagattgatttaaaacttcatttttaatttaaaaggatc

taggtgaagatccttttgataatctcatgacaaaaatccctaacgtgagtttcttccactgagcgtcagaccc
 cgtagaaaagatcaaaggatcttcttgagatcctttttctgcgcgtaatctgctgcttgcacacaaaaaaccac
 cgctaccagcgggtggttgttgcggatcaagagctaccaactcttttccgaaggtaactggcttcagcagagcg
 cagataccaaatactgtccttctagtgtagccgtagttaggccaccacttcaagaactctgtagcaccgcctacata
 cctcgctctgctaactcgttaccagtggctgctgccaagtggcgataagtcgtgtcttaccgggttgactcaagac
 gatagttaccggataaggcgcagcggctcgggctgaacggggggttcgtgcacacagcccagcttgagcgaacgacc
 tacaccgaactgagatacctacagcgtgagctatgagaaagcgccacgttcccgaaggagaaaggcggacaggta
 tccggtaaagcggcagggtcggaaacaggagagcgcacgaggagcttccagggggaacgcctggtatctttatagtc
 ctgtcgggttccgacacctgacttgagcgtcgattttgtgatgctcgtcagggggcgagcctatggaaaaac
 gccagcaacgcggccttttacggttcctggccttttctggccttttctcacatgt

**pDO242 (SaCas9 used in all JCR89/91 projects and JCR157/160 projects for in vitro work;
 SaCas9 in uppercase)(SEQ ID NO: 83)**

ctaaattgaagcgttaatatatttgttaaaattcgcgttaaattttgttaaatcagctcatttttaaccaatagg
 ccgaaatcggcaaaatccctataaatcaaaagaatagaccgagatagggttgagtgttgcagtttgaacaag
 agtccactattaaagaacgtggactccaacgtcaaaggcgaaaaaccgtctatcaggcgatggccactacgtga
 accatcacctaatacaagtttttggggtcgaggtgccgtaaagcactaaatcggaaccctaaggagccccgat
 ttagagcttgacggggaaagccggcgaaactggcgagaaaggaagggaagaaagcgaaggagcggcgctaggcg
 ctggcaagtgtagcggtcacgtgcgcgttaaccaccacaccgcccgcgttaatgcgccgtacagggcgctccca
 ttgccattcaggctgcgcaactgttgggaaggcgatcgggtcgggcctcttcgctattacgccagctggcgaaag
 ggggatgtgctgcaaggcgattaagtgggtaacgccagggtttccagtcacgacgttgtaaacgacggccagt
 gagcgcgcgtaatacactactatagggcgaattgggtacCtttaattctagtactatgcaTgcgttgacattgat
 tattgactagtattataatagtaataattacggggctcattagttcatagcccatatatggagtccgcgttacataa
 ctacggtaaatggcccgcctggctgaccgcccacgacccccgccattgacgtcaataatgacgtatgttcccat
 agtaacgccaatagggaactttcattgacgtcaatgggtggagtatttacggtaaaactgccacttggcagtacatc
 aagtgtatcatatgcaagtagccccctattgacgtcaatgacggtaaatggccgcctggcattatgccagtag
 atgacctatgggactttcctacttggcagtagatctacgtattagtcacgtattaccatgggtgatgcgggtttg
 gcagtacatcaatgggcgtggatagcgggttgactcacggggatttccaagtctccacccattgacgtcaatggga
 gttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgcccattgacgcaaatgggcggt
 aggcgtgtacggtgggaggtctatataagcagagctctctggctaactaccggtgccaccATGAAAAGGAACTACAT
 TCTGGGGCTGGACATCGGGATTACAAGCGTGGGGTATGGGATTATTGACTATGAAA

CAAGGGACGTGATCGACGCAGGCGTCAGACTGTTCAAGGAGGCCAACGTGGAAAAC
AATGAGGGACGGAGAAGCAAGAGGGGAGCCAGGCGCCTGAAACGACGGAGAAGGC
ACAGAATCCAGAGGGTGAAGAACTGCTGTTTCGATTACAACCTGCTGACCGACCAT
TCTGAGCTGAGTGGAATTAATCCTTATGAAGCCAGGGTGAAAGGCCTGAGTCAGAA
GCTGTCAGAGGAAGAGTTTTCCGCAGCTCTGCTGCACCTGGCTAAGCGCCGAGGAGT
GCATAACGTCAATGAGGTGGAAGAGGACACCGGCAACGAGCTGTCTACAAAGGAAC
AGATCTCACGCAATAGCAAAGCTCTGGAAGAGAAGTATGTCGCAGAGCTGCAGCTG
GAACGGCTGAAGAAAGATGGCGAGGTGAGAGGGTCAATTAATAGGTTCAAGACAA
GCGACTACGTCAAAGAAGCCAAGCAGCTGCTGAAAGTGCAGAAGGCTTACCACCAG
CTGGATCAGAGCTTCATCGATACTTATATCGACCTGCTGGAGACTCGGAGAACCTAC
ATGAGGGACCAGGAGAAGGGAGCCCCTTCGGATGGAAAGACATCAAGGAATGGTA
CGAGATGCTGATGGGACATTGCACCTATTTTCCAGAAGAGCTGAGAAGCGTCAAGT
ACGCTTATAACGCAGATCTGTACAACGCCCTGAATGACCTGAACAACCTGGTCATCA
CCAGGGATGAAAACGAGAACTGGAATACTATGAGAAGTTCCAGATCATCGAAAAC
GTGTTTAAGCAGAAGAAAAAGCCTACACTGAAACAGATTGCTAAGGAGATCCTGGT
CAACGAAGAGGACATCAAGGGCTACCGGGTGACAAGCACTGGAAAACCAGAGTTC
ACCAATCTGAAAGTGTATCACGATATTAAGGACATCACAGCACGGAAAGAATCATT
GAGAACGCCGAACCTGCTGGATCAGATTGCTAAGATCCTGACTATCTACCAGAGCTCC
GAGGACATCCAGGAAGAGCTGACTAACCTGAACAGCGAGCTGACCCAGGAAGAGA
TCGAACAGATTAGTAATCTGAAGGGGTACACCGGAACACACAACCTGTCCCTGAAA
GCTATCAATCTGATTCTGGATGAGCTGTGGCATACAAACGACAATCAGATTGCAATC
TTTAACCGGCTGAAGCTGGTCCCAAAAAAGGTGGACCTGAGTCAGCAGAAAGAGAT
CCCAACCACACTGGTGGACGATTTCAATTCTGTCACCCGTGGTCAAGCGGAGCTTCAT
CCAGAGCATCAAAGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAATGATA
TCATTATCGAGCTGGCTAGGGAGAAGAACAGCAAGGACGCACAGAAGATGATCAAT
GAGATGCAGAAACGAAACCGGCAGACCAATGAACGCATTGAAGAGATTATCCGAAC
TACCGGGAAAGAGAACGCAAAGTACCTGATTGAAAAAATCAAGCTGCACGATATGC
AGGAGGGAAAGTGTCTGTATTCTCTGGAGGCCATCCCCCTGGAGGACCTGCTGAAC
AATCCATTCAACTACGAGGTCGATCATATTATCCCCAGAAGCGTGTCTTCGACAAT
TCCTTTAACAACAAGGTGCTGGTCAAGCAGGAAGAGAACTCTAAAAAGGGCAATAG
GACTCCTTTCCAGTACCTGTCTAGTTCAGATTCCAAGATCTCTTACGAAACCTTTAAA

AAGCACATTCTGAATCTGGCCAAAGGAAAGGGCCGCATCAGCAAGACCAAAAAGG
AGTACCTGCTGGAAGAGCGGGACATCAACAGATTCTCCGTCCAGAAGGATTTTATTA
ACCGGAATCTGGTGGACACAAGATACGCTACTCGCGGCTGATGAATCTGCTGCGAT
CCTATTTCCGGGTGAACAATCTGGATGTGAAAGTCAAGTCCATCAACGGCGGGTTCA
CATCTTTTCTGAGGCGCAAATGGAAGTTTAAAAAGGAGCGCAACAAAGGGTACAAG
CACCATGCCGAAGATGCTCTGATTATCGCAAATGCCGACTTCATCTTTAAGGAGTGG
AAAAAGCTGGACAAAGCCAAGAAAGTGATGGAGAACCAGATGTTCTGAAGAGAAGC
AGGCCGAATCTATGCCCCGAAATCGAGACAGAACAGGAGTACAAGGAGATTTTCATC
ACTCCTCACCAGATCAAGCATATCAAGGATTTCAAGGACTACAAGTACTCTCACCGG
GTGGATAAAAAGCCCAACAGAGAGCTGATCAATGACACCCTGTATAGTACAAGAAA
AGACGATAAGGGGAATACCCTGATTGTGAACAATCTGAACGGACTGTACGACAAAG
ATAATGACAAGCTGAAAAAGCTGATCAACAAAAGTCCCGAGAAGCTGCTGATGTAC
CACCATGATCCTCAGACATATCAGAAACTGAAGCTGATTATGGAGCAGTACGGCGA
CGAGAAGAACCCACTGTATAAGTACTATGAAGAGACTGGGAACTACCTGACCAAGT
ATAGCAAAAAGGATAATGGCCCCGTGATCAAGAAGATCAAGTACTATGGGAACAAG
CTGAATGCCCATCTGGACATCACAGACGATTACCCTAACAGTCGCAACAAGGTGGT
CAAGCTGTCACTGAAGCCATACAGATTCGATGTCTATCTGGACAACGGCGTGTATAA
ATTTGTGACTGTCAAGAATCTGGATGTCATCAAAAAGGAGAACTACTATGAAGTGA
ATAGCAAGTGCTACGAAGAGGCTAAAAAGCTGAAAAAGATTAGCAACCAGGCAGA
GTTTCATCGCCTCCTTTTACAACAACGACCTGATTAAGATCAATGGCGAACTGTATAG
GGTCATCGGGGTGAACAATGATCTGCTGAACCGCATTGAAGTGAATATGATTGACAT
CACTTACCGAGAGTATCTGGAAAACATGAATGATAAGCGCCCCCTCGAATTATCAA
AACAATTGCCTCTAAGACTCAGAGTATCAAAAAGTACTCAACCGACATTCTGGGAA
ACCTGTATGAGGTGAAGAGCAAAAAGCACCCCTCAGATTATCAAAAAGGGCagcgagggc
aagcgt cctgctgctactaagaaagctggtaagctaagaaaaagaaaggatcctaccatac gatgttcagattacgttta
agaattcctagagctcgctgatcagcctcgactgtgccttctagtgtccagccatctgtgtttgccccctccccgt
gccttccttgaccctggaagggtgccactcccactgtcctttcctaataaaatgaggaaattgcatgcgattgtctga
gtagggtgcattctattctgggggggtgggggtggggcaggacagcaagggggaggattgggaagagaatagcaggcat
gctggggaggtagcgggccgcCCgcggtggagctccagcttttgtcccttagtgagggttaattgcgcgcttggcg
taatcatggatcatagctgtttcctgtgtgaaattgttatccgctcacaaattccacacaacatacgagccggaagcat
aaagtgtaaagcctgggggtgcctaattgagtgaagtaactcacattaattgcgttgcgctcactgccccgtttccagt

pJRH1 (SaCas9 used for all JCR179/183 projects, SaCas9 is in uppercase; NLS is lowercase, bolded, and underlined)(SEQ ID NO: 84)

- 106C-

agtccactattaaagaacgtggactccaacgtcaaagggcgaaaaaccgtctatcagggcgatggcccactacgtga
accatcacctaatacaagtttttggggctgaggtgccgtaaagcactaaatcggaaccctaaagggagccccgat
ttagagcttgacggggaaagccggcgaaacgtggcgagaaaggaagggaagaaagcgaaaggagcgggcgctagggcg
ctggcaagtgtagcgggtcacgtgcgcgtaaccaccacacccgccgcgttaatgcgccgtacagggcgcgctccca
ttcgccattcaggctgcgcaactgttgggaagggcgatcggtgcgggcctcttcgctattacgccagctggcgaaag
ggggatgtgctgcaaggcgattaagttgggtaacgccagggttttcccagtcacgacgttgtaaacgacggccagt
gagcgcgcgtaatacgaactactatagggcgaaattgggtaccttaattctagtactatgcatgcgttgacattgat
tattgactagttattaatagtaatacaattacggggtcattagttcatagcccatatatggagttccgcgttacataa
cttacggtaaatggcccgcctgggtgaccgccaacgacccccgccattgacgtcaataatgacgtatgttcccat
agtaacgccaatagggactttccattgacgtcaatgggtggagtatttacggtaaacgccacttggcagtacatc
aagtgtatcatatgcaagtagccccctattgacgtcaatgacggtaaatggcccgcctggcattatgccagtac
atgacctatgggactttcctacttggcagtagacatctacgtattatgcatcgctattaccatggtgatgcggttttg
gcagtacatcaatgggcgtggatagcggtttgactcacggggatttccaagtctccacccattgacgtcaatggga
gtttgttttggcaccaaaatcaacgggactttccaaaatgtcgtaacaactccgcccattgacgcaaatgggcggt
aggcgtgtacgggtgggaggtctatataagcagagctctctggctaactaccggtgccaccatggcccaaaagaagaa
gcggaaggtcggtatccacggagtcccagcagccAAGCGGAACCTACATCCTGGGCCTGGACATCGGCA
TCACCAGCGTGGGCTACGGCATCATCGACTACGAGACACGGGACGTGATCGATGCC
GGCGTGCGGCTGTTCAAAGAGGCCAACGTGGAAAACAACGAGGGCAGGCGGAGCA
AGAGAGGCGCCAGAAGGCTGAAGCGGCGGAGGCGGCATAGAATCCAGAGAGTGAA
GAAGCTGCTGTTCGACTACAACCTGCTGACCGACCACAGCGAGCTGAGCGGCATCA
ACCCCTACGAGGCCAGAGTGAAGGGCCTGAGCCAGAAGCTGAGCGAGGAAGAGTT
CTCTGCCGCCCTGCTGCACCTGGCCAAGAGAAGAGGCGTGCACAACGTGAACGAGG
TGGAAGAGGACACCGGCAACGAGCTGTCCACCAAGAGCAGATCAGCCGGAACAG
CAAGGCCCTGGAAGAGAAATACGTGGCCGAACCTGCAGCTGGAACGGCTGAAGAAA
GACGGCGAAGTGCGGGGCAGCATCAACAGATTCAAGACCAGCGACTACGTGAAAG
AAGCCAAACAGCTGCTGAAGGTGCAGAAGGCCTACCACCAGCTGGACCAGAGCTTC
ATCGACACCTACATCGACCTGCTGGAAACCCGGCGGACCTACTATGAGGGACCTGG
CGAGGGCAGCCCCTTCGGCTGGAAGGACATCAAAGAATGGTACGAGATGCTGATGG
GCCACTGCACCTACTTCCCCGAGGAACTGCGGAGCGTGAAGTACGCCTACAACGCC
GACCTGTACAACGCCCTGAACGACCTGAACAATCTCGTGATCACCAGGGACGAGAA
CGAGAAGCTGGAATATTACGAGAAGTTCCAGATCATCGAGAACGTGTTCAAGCAGA

AGAAGAAGCCCCACCCTGAAGCAGATCGCCAAAGAAATCCTCGTGAACGAAGAGGA
TATTAAGGGCTACAGAGTGACCAGCACCGGCAAGCCCGAGTTCACCAACCTGAAGG
TGTACCACGACATCAAGGACATTACCGCCCCGAAAGAGATTATTGAGAACGCCGAG
CTGCTGGATCAGATTGCCAAGATCCTGACCATCTACCAGAGCAGCGAGGACATCCA
GGAAGAACTGACCAATCTGAACTCCGAGCTGACCCAGGAAGAGATCGAGCAGATCT
CTAATCTGAAGGGCTATAACCGGCACCCACAACCTGAGCCTGAAGGCCATCAACCTG
ATCCTGGACGAGCTGTGGCACACCAACGACAACCAGATCGCTATCTTCAACCGGCT
GAAGCTGGTGCCCAAGAAGGTGGACCTGTCCCAGCAGAAAGAGATCCCCACCACCC
TGGTGGACGACTTCATCCTGAGCCCCGTCGTGAAGAGAAGCTTCATCCAGAGCATCA
AAGTGATCAACGCCATCATCAAGAAGTACGGCCTGCCCAACGACATCATTATCGAG
CTGGCCCGCGAGAAGAAGTCCAAGGACGCCAGAAAATGATCAACGAGATGCAGA
AGCGGAACCGGCAGACCAACGAGCGGATCGAGGAAATCATCCGGACCACCGGCAA
AGAGAACGCCAAGTACCTGATCGAGAAGATCAAGCTGCACGACATGCAGGAAGGC
AAGTGCCTGTACAGCCTGGAAGCCATCCCTCTGGAAGATCTGCTGAACAACCCCTTC
AACTATGAGGTGGACCACATCATCCCCAGAAGCGTGTCCCTTCGACAACAGCTTCAAC
AACAAGGTGCTCGTGAAGCAGGAAGAAAACAGCAAGAAGGGGCAACCGGACCCCAT
TCCAGTACCTGAGCAGCAGCGACAGCAAGATCAGCTACGAAACCTTCAAGAAGCAC
ATCCTGAATCTGGCCAAGGGCAAGGGCAGAATCAGCAAGACCAAGAAAGAGTATCT
GCTGGAAGAACGGGACATCAACAGGTTCTCCGTGCAGAAAGACTTCATCAACCGGA
ACCTGGTGGATACCAGATACGCCACCAGAGGCCTGATGAACCTGCTGCGGAGCTAC
TTCAGAGTGAACAACCTGGACGTGAAAGTGAAGTCCATCAATGGCGGCTTCAACAG
CTTTCTGCGGCGGAAGTGGAAGTTTAAGAAAGAGCGGAACAAGGGGTACAAGCACC
ACGCCGAGGACGCCCTGATCATTGCCAACGCCGATTTTCATCTTCAAAGAGTGGAAG
AAACTGGACAAGGCCAAAAAAGTGATGGAAAACCAGATGTTCGAGGAAAAGCAGG
CCGAGAGCATGCCCGAGATCGAAACCGAGCAGGAGTACAAAGAGATCTTCATCACC
CCCCACCAGATCAAGCACATTAAGGACTTCAAGGACTACAAGTACAGCCACCGGGT
GGACAAGAAGCCTAATAGAGAGCTGATTAACGACACCCTGTACTCCACCCGGAAGG
ACGACAAGGGCAACACCCTGATCGTGAACAATCTGAACGGCCTGTACGACAAGGAC
AATGACAAGCTGAAAAAGCTGATCAACAAGAGCCCCGAAAAGCTGCTGATGTACCA
CCACGACCCCCAGACCTACCAGAACTGAAGCTGATTATGGAACAGTACGGCGACG
AGAAGAATCCCCTGTACAAGTACTACGAGGAAACCGGGAACCTGACCAAGTAC

TCCAAAAAGGACAACGGCCCCGTGATCAAGAAGATTAAGTATTACGGCAACAACT
GAACGCCCATCTGGACATCACCGACGACTACCCCAACAGCAGAAACAAGGTCGTGA
AGCTGTCCCTGAAGCCCTACAGATTGACGTGTACCTGGACAATGGCGTGTACAAGT
TCGTGACCGTGAAGAATCTGGATGTGATCAAAAAAGAAAATACTACGAAGTGAAT
AGCAAGTGCTATGAGGAAGCTAAGAAGCTGAAGAAGATCAGCAACCAGGCCGAGTT
TATCGCCTCCTTCTACAACAACGATCTGATCAAGATCAACGGCGAGCTGTATAGAGT
GATCGGCGTGAACAACGACCTGCTGAACCGGATCGAAGTGAACATGATCGACATCA
CCTACCGCGAGTACCTGGAAAACATGAACGACAAGAGGCCCCCAGGATCATTAAAG
ACAATCGCCTCCAAGACCCAGAGCATTAAAGAAGTACAGCACAGACATTCTGGGCAA
CCTGTATGAAGTGAAATCTAAGAAGCACCCCTCAGATCATCAAAAAGGGCaaaaggccgg
cggccacgaaaaaggccggccaggcaaaaaagaa

aaagggatcctaccatacgaatgtccagattacgcttaagaattcctagagctcgtgatcagcctcgactgtgcc
ttctagtggcagccatctgtgtttgccccccccctgccttccttgaccctggaaggtgccactcccactgtcc
ttcctaataaaatgaggaaattgcatcgcattgtctgagtaggtgtcattctattctgggggggtgggggtggggcag
gacagcaagggggaggattgggaagagaatagcaggcatgctggggaggtagcggccgcccgcggtggagctccagc
tttgttcccttagtgagggttaattgcgcgcttggcgtaatcatggcatagctgttccctgtgtgaaattgtta
tcgctcacaattccacacaacatacagcgggaagcataaagtgtaaagcctgggggtgcctaatgagtgcgtaac
tcacattaattgcgttgcgctcactgcccgtttccagtcgggaaacctgtcgtgccagctgcattaatgaatcggc
caacgcgcggggagaggcggttgcgtattggcgctcttccgttctcgtcactgactcgtcgcgtcggctgt
tcggctgcggcgagcggatcagctcactcaaaggcggtatccaggtatccacagaatcaggggataacgcaggaa
agaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgcgttgctggcggttttccataggctc
cgccccctgacgagcatcacaaaaatcgacgtcaagtcagaggtggcgaaacccgacaggactataagatacca
ggcggttccccctggaagctccctcgtgcgctctcctgttccgacctgccgcttaccggatacctgtccgccttc
tcccttcgggaagcgtggcgctttctcatagctcacgctgtaggtatctcagttcgggtgtaggtcgttcgtccaag
ctgggctgtgtgcacgaacccccgttcagcccagccgctgcgccttatccggtactatcgtcttgagccaaccc
ggtaagacacgacttatgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcgggtgcta
cagagttctgaagtgggtggcctaactacggctacactagaaggacagtatttggtatctgcgctcgtcgaagcca
gttaccttcgaaaaagagttggtagctcttgatccggcaaaacaccacgctggtagcgggtgtttttgtttg
caagcagcagattacgcgcagaaaaaaggatctcaagaagatccttgatctttctacggggtcgtacgctcagt
ggaacgaaaactcacgttaagggttttggctcatgagattatcaaaaaggatcttcacctagatcctttaaattaa
aaatgaagtttaaatcaatctaaagtatatatgagtaaacttggctgcaggtaccaatgcttaatcagtgaggc

acctatctcagcgatctgtctatttcgttcatccatagttgcctgactccccgtcgtgtagataactacgatacggg
agggcttaccatctggccccagtgctgcaatgataccgcgagaccacgctcaccggctccagatttatcagcaata
aaccagccagccggaaggccgagcgcagaagtggctctgcaactttatccgctccatccagtcatttaattgttg
ccgggaagctagagtaagtagttcgccagttaatagtttgcgcaacgttggtgccattgctacaggcatcgtggtgt
cacgctcgtcgttttggtatggcttcattcagctccgggtcccaacgatcaaggcgagttacatgatccccatgttg
tgcaaaaaagcgggttagctccttcgggtcctccgatcgttgtcagaagtaagttggccgcagtggtatcactcatggt
tatggcagcactgcataattcttactgtcatgccatccgtaagatgcttttctgtgactgggtgagtactcaacca
agtcattctgagaatagtgtatgcggcgaccgagttgctcttgcggcgcaatacgggataataccgcgccacat
agcagaactttaaaagtgtcatcattggaaaacgttcttcggggcgaaaactctcaaggatcttaccgctgttgag
atccagttcgatgtaaccactcgtgcaccaactgatcttcagcatcttttactttaccagcgtttctgggtgag
caaaaacaggaaggcaaaatgccgcaaaaaagggaataaggcgacacggaaatgtgaatactcatactcttctt
ttcaatattattgaagcatttatcaggggtattgtctcatgagcggatacatattgaatgtattagaaaaataa
acaaataggggtccgcgcacatttccccgaaaagtgccac

NLS sequence in PT366 (SEQ ID NO: 85)

AAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCACAAAAAGAAAAAG

pDO203 – Generic backbone to clone gRNAs into for SpCas9; JCR94 and JCR99 were put into the site in bold to test in cells then to make mRNA from for making the hDMD-delta52/mdx mouse (SEQ ID NO: 86)

ctaaattgtaagcgtaatatattttgttaaaattcgcgttaaattttgttaaatcagctcatttttaaccaatagg
ccgaaatcggcataatccctataaatcaaaagaatagaccgagataggggtgagtggttccagtttgaacaag
agtccactattaaagaacgtggactccaacgtcaaaaggcgaaaaaccgtctatcaggcgcatggcccactacgtga
accatcacctaatacaagtttttggggtcgaggtgccgtaaagcactaaatcggaaccctaaaggagccccgat
ttagagcttgacggggaaagccggcgaaacgtggcgagaaagggaaggaaagcgaaggagcggcgctagggcg
ctggcaagtgtagcggtcacgtgcgcgtaaccaccacaccgccgcgttaatgcgccgtacaggcgcgctcca
ttcgccattcaggctgcgcaactgttgggaaggcgatcgggtcgggcctcttcgctattacgccagctggcgaaag
ggggatgtctgcaaggcgattaagttgggtaacccagggtttccagtcacgacgttgtaaacgacggccagt
gagcgcgcgtaatacgaactcactatagggcgaattgggtaccgagggcctatttccatgattccttcatatttga
tatacgatacaaggctgttagagagataattggaattaatttgactgtaaacacaaagatattagtaaaaatcgt
gacgtagaaagtaataatttctgggtagtttcagttttaaaattatgtttaaaatggactatcatatgcttacc

gtaacttgaaagtatttcgatttcttggctttatatatcttgtggaaggacgaaacaccGGGTCTTCGAGAAGACC
Tgttttagagctagaaatagcaagttaaaataaggctagtcggtatcaacttgaaaaagtgccaccgagtcggtgc
ttttttccgcggtggagctccagcttttgtcccttagtgagggttaattgcgcgcttggcgtaatcatggtcat
agctgtttcctgtgtgaaattgttatccgtcacaattccacacaacatacagccggaagcataaagtgtaaagcc
tggggtgcctaataagtgagctaacacattaattgcgttgcgtcactgcccgtttcagtcgggaaacctgtc
gtgccagctgcattaatgaatcggccaacgcgcggggagaggcggttgcgtattgggcgctcttccgcttctcgc
tactgactcgtcgcgtcggctgttcggctgcggcgagcgggtatcagctcactcaaaggcggtaatcaggttatcc
acagaatcaggggataacgcaggaaagaacatgtgagcaaaaggccagcaaaaggccaggaaccgtaaaaaggccgc
gttgcgtggcggttttccataggctccgccccctgacgagcatcacaaaaatcgacgctcaagtcagaggtggcgaa
accgcagaggactataaagataaccaggcggttccccctggaagctccctcgtgcgctctctgttccgacctgccc
cttaccggatacctgtccgcctttctcccttgggaagcgtggcgctttctcatagctcacgctgtaggtatctcag
ttcgggtgtaggtcgttcgctccaagctgggctgtgtgcacgaacccccgttcagccccaccgctgcgccttatccg
gtaactatcgtcttgagtccaaccggtaagacacgacttatcgccactggcagcagccactggtaacaggattagc
agagcgaggtatgtaggcgggtgtacagagtcttgaagtgggtggcctaactacggctacactagaaggacagtatt
tggtatctgcgctctgctgaagccagttaccttcgaaaaagagttggtagctcttgatccggcaaaaaaccaccg
ctggtagcggtgggtttttgttgaagcagcagattacgcgcagaaaaaaaggatctcaagaagatcctttgatc
ttttctacgggggtctgacgctcagtggaacgaaaactcacgttaagggttttggcatgagattatcaaaaaggat
cttcacctagatccttttaattaaaaatgaagttttaaataaatctaaagtatatatgagtaaaacttggctgaca
gttaccaatgcttaatacagtgaggcacctatctcagcgatctgtctatttcgttcacatagttgctgactcccc
gtcgtgtagataactacgatacgggagggcttaccatctggccccagtgctgcaatgataccgcgagaccacgctc
accggctccagattatcagcaataaaccagccagccggaaggggcgagcgcagaagtggctctgcaactttatccg
ctccatccagctatttaattgttgcgggaagctagagtaagtagttccgagtaataagtttgcgcaacgttgtt
gccattgctacaggcatcgtggtgtcacgctcgtcttgggtatggcttcattcagctccggttccaacgatcaag
gcgagttacatgatccccatgttgtgcaaaaaagcggttagctccttcggctcctccgatcggtgtcagaagtaagt
tgccgcagtggtatcactcatggttatggcagcactgcataattctcttactgtcatgccatccgtaagatgttt
tctgtgactgggtgagtactcaaccaagtcattctgagaatagtgatgcggcgaccgagttgctcttggccggcgtc
aatacgggataataccgcgccacatagcagaactttaaaagtgtcatcattggaaaacgttcttcggggcgaaaaa
tctcaaggatcttaccgctgttgagatccagttcagtgtaaccactcgtgcaccaactgatcttcagcatctttt
actttcaccagcggttctgggtgagcaaaaacaggaaaggcaaaatgccgcaaaaaagggaataaggggcgacacggaa
atgttgaatactcatactcttcttttcaatattattgaagcatttatcagggttattgtctcatgagcggataca
tattgaatgtatttagaaaaataacaaataggggttccgcgcacatttccccgaaaagtgccac

CLAIMS

What is claimed is:

1. A guide RNA (gRNA) comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42 or a complement thereof.
2. A DNA targeting composition comprising a first gRNA and a second gRNA, the first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42, or a complement thereof, wherein the first gRNA molecule and the second gRNA molecule comprise different targeting domains.
3. The DNA targeting composition claim 2, wherein the first gRNA molecule is SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, or SEQ ID NO: 41, and the second gRNA molecule is SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, or SEQ ID NO: 42.
4. The DNA targeting composition of claim 2, wherein the first gRNA molecule is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14 and SEQ ID NO: 15, and the second gRNA molecule is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, and SEQ ID NO: 19.

5. The DNA targeting composition of claim 2, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of:

(i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2;

(ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(viii) a first gRNA molecule comprising a targeting domain that comprises a

nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15;

(xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18; and

(xii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 41, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 42.

6. The DNA targeting composition of claim 2, further comprising a Clustered Regularly Interspaced Short Palindromic Repeats associated (Cas) protein.

7. The DNA targeting composition of claim 6, wherein the Cas protein comprises a Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25).

8. The DNA targeting composition of claim 6, wherein the Cas protein comprises a *Staphylococcus aureus* Cas9 molecule having an amino acid sequence of SEQ ID NO: 45.

9. The DNA targeting composition of claim 2, wherein the DNA targeting composition comprises a nucleotide sequence of SEQ ID NO: 83, a nucleotide sequence of SEQ ID NO: 84, a nucleotide sequence of SEQ ID NO: 37, and/or a nucleotide sequence of SEQ ID NO: 38.

10. An isolated polynucleotide comprising the gRNA molecule of claim 1.
11. A vector comprising the gRNA of claim 1.
12. A vector comprising the DNA targeting composition of claim 6.
13. A vector encoding:
 - (a) a first guide RNA (gRNA) molecule,
 - (b) a second gRNA molecule, and
 - (c) at least one Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25),wherein the first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 41, SEQ ID NO: 42, or a complement thereof, wherein the first gRNA molecule and the second gRNA molecule comprise different targeting domains.
14. The vector of claim 13, wherein the vector is configured to form a first and a second double strand break in a first and a second intron flanking exon 51 of the human *DMD* gene.
15. The vector of claim 13, wherein the first gRNA molecule is SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, or SEQ ID NO: 41, and the second gRNA molecule is SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, or SEQ ID NO: 42.
16. The vector of claim 13, wherein the first gRNA molecule is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14 and SEQ ID NO: 15, and the second gRNA molecule is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, and SEQ ID NO: 19.

17. The vector of claim 13, the first gRNA molecule and the second gRNA molecule are selected from the group consisting of:

(i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2;

(ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(viii) a first gRNA molecule comprising a targeting domain that comprises a

nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; and

(xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18.

18. The vector of claim 11, wherein the vector is a viral vector.

19. The vector of claim 18, wherein the vector is an Adeno-associated virus (AAV) vector.

20. The vector of claim 19, wherein the AAV vector is an AAV8 vector or an AAV9 vector.

21. The vector of claim 11, wherein the vector comprises a tissue-specific promoter operably linked to the nucleotide sequence encoding the first gRNA molecule, the second gRNA molecule, and/or the Cas9 molecule.

22. The vector of claim 21, wherein the tissue-specific promoter is a muscle specific promoter.

23. A cell comprising the gRNA of claim 1.

24. A kit comprising the gRNA of claim 1.

25. A method of correcting a mutant *dystrophin* gene in a cell, the method comprising administering to a cell the gRNA of claim 1.

26. A method of genome editing a mutant *dystrophin* gene in a subject, the method comprising administering to the subject a genome editing composition comprising the gRNA of claim 1.

27. The method of claim 26, wherein the genome editing composition is administered to the subject intramuscularly, intravenously, or a combination thereof.

28. The method of claim 25 or 26, wherein correcting the mutant *dystrophin* gene comprises nuclease-mediated non-homologous end joining.

29. A method of treating a subject in need thereof having a mutant *dystrophin* gene, the method comprising administering to the subject the gRNA of claim 1.

30. A modified adeno-associated viral vector for genome editing a mutant *dystrophin* gene in a subject comprising a first polynucleotide sequence encoding the gRNA of claim 1, and a second polynucleotide sequence encoding a Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25).

31. The modified adeno-associated viral vector of claim 30, wherein the modified adeno-associated viral vector comprises the nucleotide sequence set forth in SEQ ID NO: 39 or SEQ ID NO: 40.

32. A composition for deleting a segment of a *dystrophin* gene comprising exon 51, the composition comprising:

(a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and

(b) a second vector comprising a polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25),

wherein each of the first and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the first vector and second vector are configured to form a first and a second double strand break in a first intron and a second intron flanking exon 51 of

the human *DMD* gene, respectively, thereby deleting a segment of the *dystrophin* gene comprising exon 51.

33. The composition of claim 32, wherein the segment has a length of about 50 base pairs to about 2,000 base pairs.

34. The composition of claim 33, wherein the segment has a length of about 118 base pairs, about 233 base pairs, about 326 base pairs, about 766 base pairs, about 805 base pairs, or about 1611 base pairs.

35. The composition of claim 32, wherein the first Cas9 molecule and the second Cas9 molecule are the same.

36. The composition of claim 35, wherein the first Cas9 molecule and the second Cas9 molecule is a *Staphylococcus aureus* Cas9 molecule.

37. The composition of claim 36, wherein the first Cas9 molecule and the second Cas9 molecule is a mutant *Staphylococcus aureus* Cas9 molecule.

38. The composition of claim 32, wherein the first Cas9 molecule and the second Cas9 molecule are different.

39. The composition of claim 38, wherein the first Cas9 molecule or the second Cas9 molecule is a *Staphylococcus aureus* Cas9 molecule.

40. The composition of claim 32, wherein the first Cas9 molecule and/or the second Cas9 molecule comprises a SaCas9 molecule having an amino acid sequence of SEQ ID NO: 45.

41. The composition of claim 32, wherein the first vector and/or the second vector is a viral vector.

42. The composition of claim 41, wherein the first vector and/or the second vector is an Adeno-associated virus (AAV) vector.

43. The composition of claim 42, wherein the AAV vector is an AAV8 vector or an AAV9 vector.

44. The composition of claim 32, wherein the *dystrophin* gene is a human *dystrophin* gene.

45. The composition of claim 32, wherein the first gRNA molecule and the second gRNA molecule comprise a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, or a complement thereof, wherein the first gRNA molecule and the second gRNA molecule comprise different targeting domains.

46. The composition of claim 32, wherein the first gRNA molecule is SEQ ID NO: 1, SEQ ID NO: 3, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14, or SEQ ID NO: 15, and the second gRNA molecule is SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, or SEQ ID NO: 19.

47. The composition of claim 32, wherein the first gRNA molecule is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 11, SEQ ID NO: 12, SEQ ID NO: 13, SEQ ID NO: 14 and SEQ ID NO: 15, and the second gRNA molecule is selected from the group consisting of SEQ ID NO: 2, SEQ ID NO: 4, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, and SEQ ID NO: 19.

48. The composition of claim 32, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of:

(i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2;

(ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID

NO: 19;

(iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; and

(xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule

comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18.

49. The composition of claim 32, wherein the first vector comprises a nucleotide sequence set forth in SEQ ID NO: 39 and the second vector comprises a nucleotide sequence set forth in SEQ ID NO: 40.

50. The composition of claim 32, for use in a medicament.

51. The composition of claim 32, for use in the treatment of Duchenne Muscular Dystrophy.

52. A cell comprising the composition of claim 32.

53. A method of correcting a mutant *dystrophin* gene in a cell, comprising administering to the cell:

(a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and

(b) a second vector comprising a polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25),

wherein each of the first gRNA and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the vector is configured to form a first and a second double strand break in a first and a second intron flanking exon 51 of the human *dystrophin* gene, respectively, thereby deleting a segment of the *dystrophin* gene comprising exon 51.

54. The method of claim 53, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of:

(i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2;

(ii) a first gRNA molecule comprising a targeting domain that comprises a

nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; and

(xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18.

55. The method of claim 53, wherein the mutant *dystrophin* gene comprises a premature stop codon, disrupted reading frame, an aberrant splice acceptor site, or an aberrant splice donor site.

56. The method of claim 53, wherein the mutant *dystrophin* gene comprises a frameshift mutation which causes a premature stop codon and a truncated gene product.

57. The method of claim 53, wherein the mutant *dystrophin* gene comprises a deletion of one or more exons which disrupts the reading frame.

58. The method of claim 53, wherein the correction of the mutant *dystrophin* gene comprises a deletion of a premature stop codon, correction of a disrupted reading frame, or modulation of splicing by disruption of a splice acceptor site or disruption of a splice donor sequence.

59. The method of claim 53, wherein the correction of the mutant *dystrophin* gene comprises deletion of exon 51.

60. The method of claim 53, wherein the correction of the mutant *dystrophin* gene comprises homology-directed repair.

61. The method of claim 60, further comprising administering to the cell a donor DNA.

62. The method of claim 53, wherein the correction of the mutant *dystrophin* gene comprises nuclease mediated non-homologous end joining.

63. The method of claim 53, wherein the cell is a myoblast cell.

64. The method of claim 53, wherein the cell is from a subject suffering from Duchenne muscular dystrophy.

65. The method of claim 53, wherein the cell is a myoblast from a human subject suffering from Duchenne muscular dystrophy.

66. The method of claim 53, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of:

(i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2;

(ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4; and

(iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19.

67. A method of treating a subject in need thereof having a mutant *dystrophin* gene, the method comprising administering to the subject:

(a) a first vector comprising a polynucleotide sequence encoding a first guide RNA (gRNA) molecule and a polynucleotide sequence encoding a first Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25), and

(b) a second vector comprising a polynucleotide sequence encoding a second gRNA molecule and a polynucleotide sequence encoding a second Cas9 molecule that recognizes a Protospacer Adjacent Motif (PAM) of either NNGRRT (SEQ ID NO: 24) or NNGRRV (SEQ ID NO: 25),

wherein each of the first gRNA and second gRNA molecules have a targeting domain of 19 to 24 nucleotides in length, and wherein the vector is configured to form a first and a second

double strand break in a first and a second intron flanking exon 51 of the human *dystrophin* gene, respectively, thereby deleting a segment of the *dystrophin* gene comprising exon 51.

68. The method of claim 67, wherein the first gRNA molecule and the second gRNA molecule are selected from the group consisting of:

(i) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 1, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 2;

(ii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(iii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(iv) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18;

(v) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(vi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(vii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID

NO: 18;

(viii) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 4;

(ix) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 19;

(x) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 14, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 15; and

(xi) a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 11, and a second gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 18.

69. The method of claim 68, wherein the subject is suffering from Duchenne muscular dystrophy.

70. The method of claim 67, wherein the first vector and second vector are administered to the subject intramuscularly, intravenously, or a combination thereof.

71. A method of generating a transgenic rodent embryo having a human *dystrophin* gene (hDMD) with an exon 52 deletion ($\Delta 52$), the method comprising administering to a rodent embryo the gRNA of claim 1.

72. The method of claim 71, wherein the rodent embryo is a mouse embryo.

73. The method of claim 71, wherein the transgenic rodent embryo is heterozygous hDMD or heterozygous hDMD- $\Delta 52$.

74. The method of claim 71, wherein a first gRNA molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 41, and a second gRNA

molecule comprising a targeting domain that comprises a nucleotide sequence set forth in SEQ ID NO: 42 are administered to the rodent embryo to delete exon 52 of the human *dystrophin* gene.

75. The method of claim 71, further comprising administering to the rodent embryo a Cas protein comprising an amino acid sequence set forth in SEQ ID NO: 27.

76. A transgenic rodent embryo produced by the method of claim 71

77. A transgenic rodent produced from the transgenic rodent embryo of claim 71.

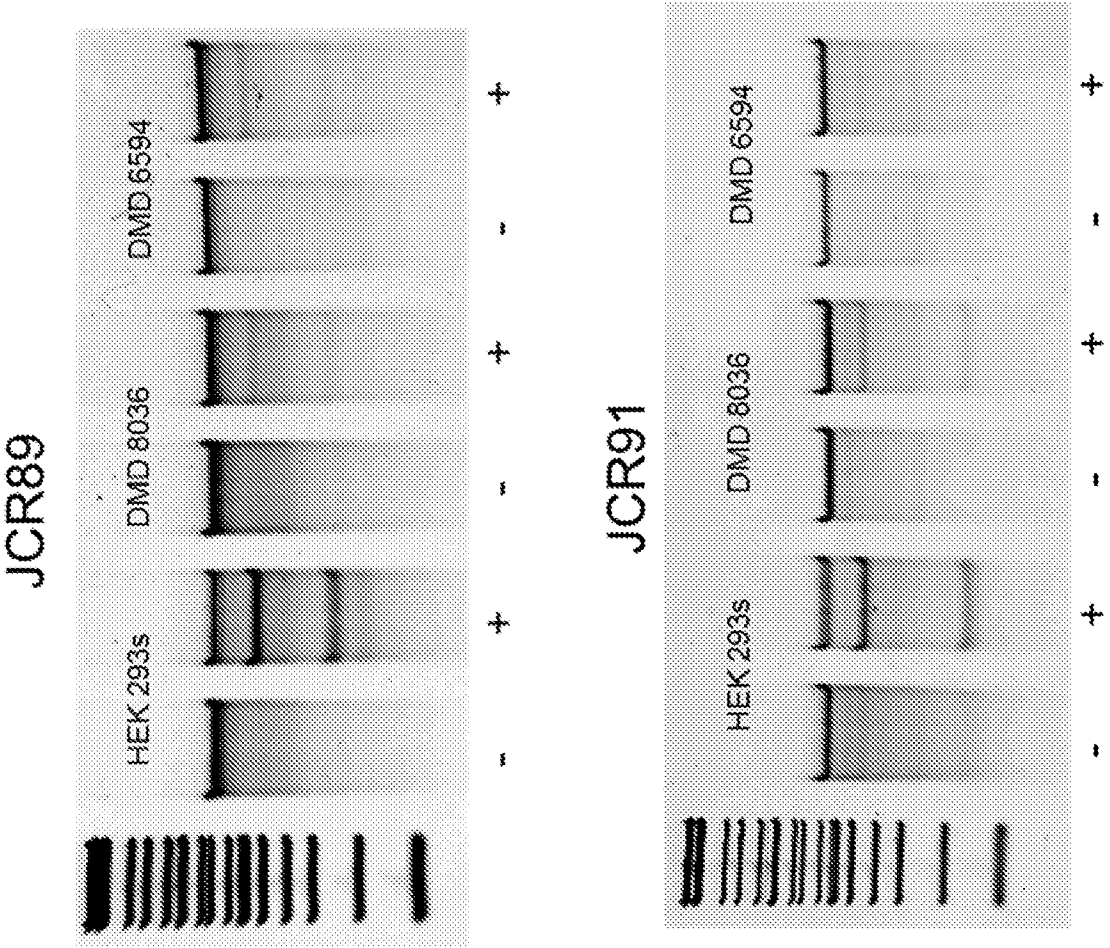


FIG. 1

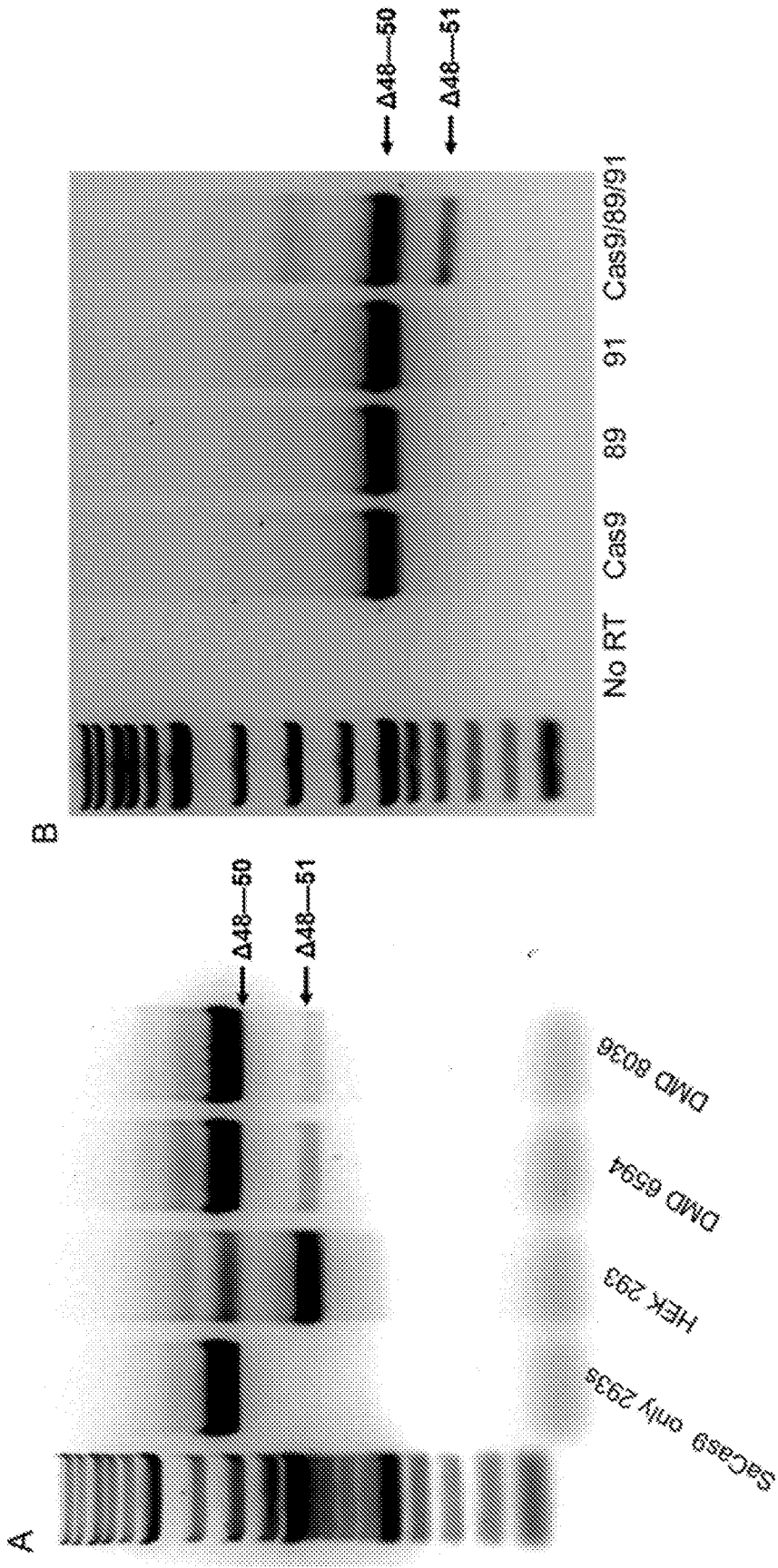
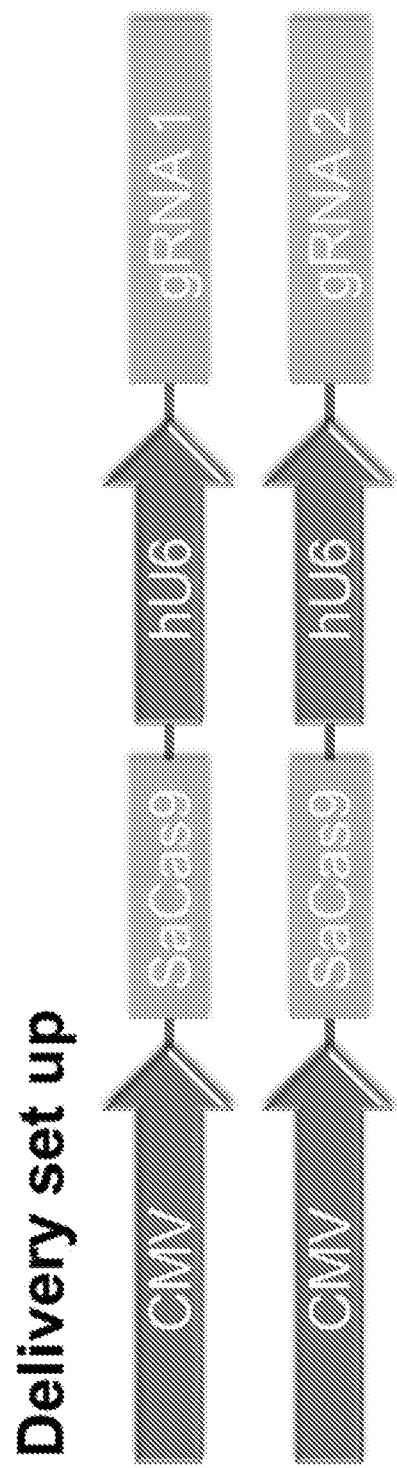


FIG. 2



Codeliver 2 viruses w/JCR89, 91

FIG. 3

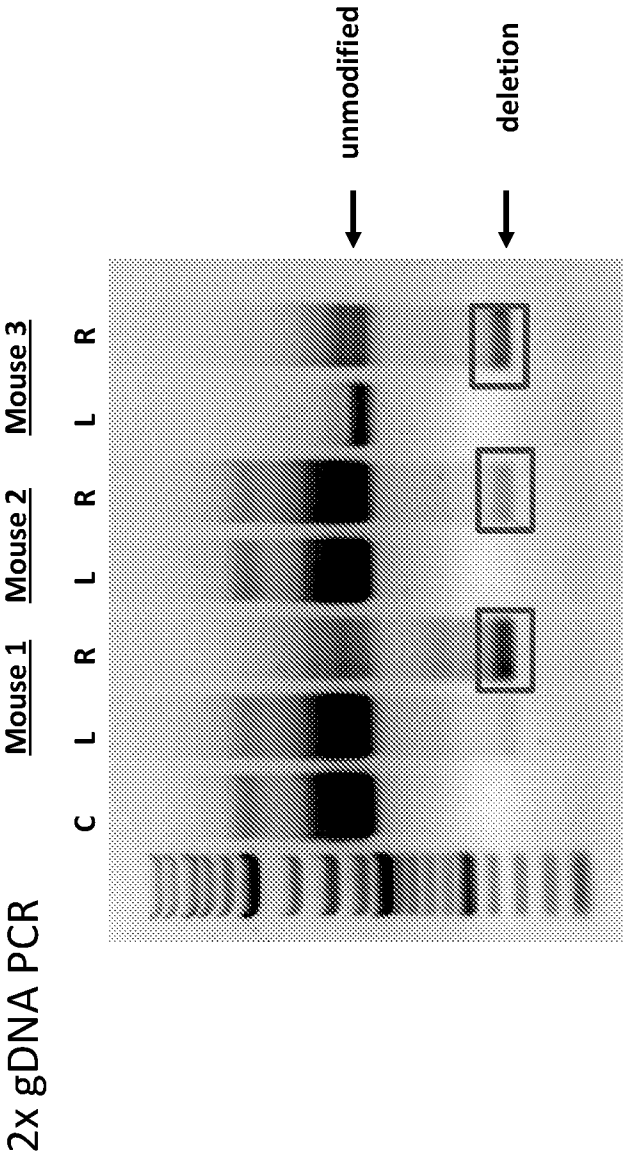


FIG. 4

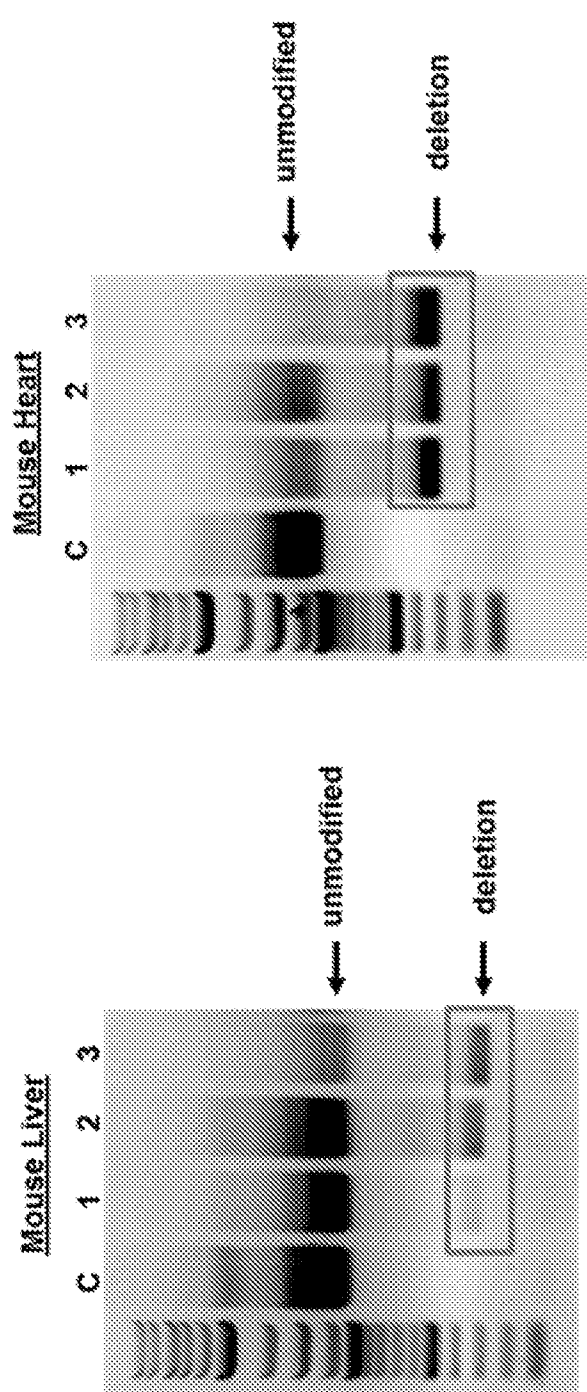


FIG. 5

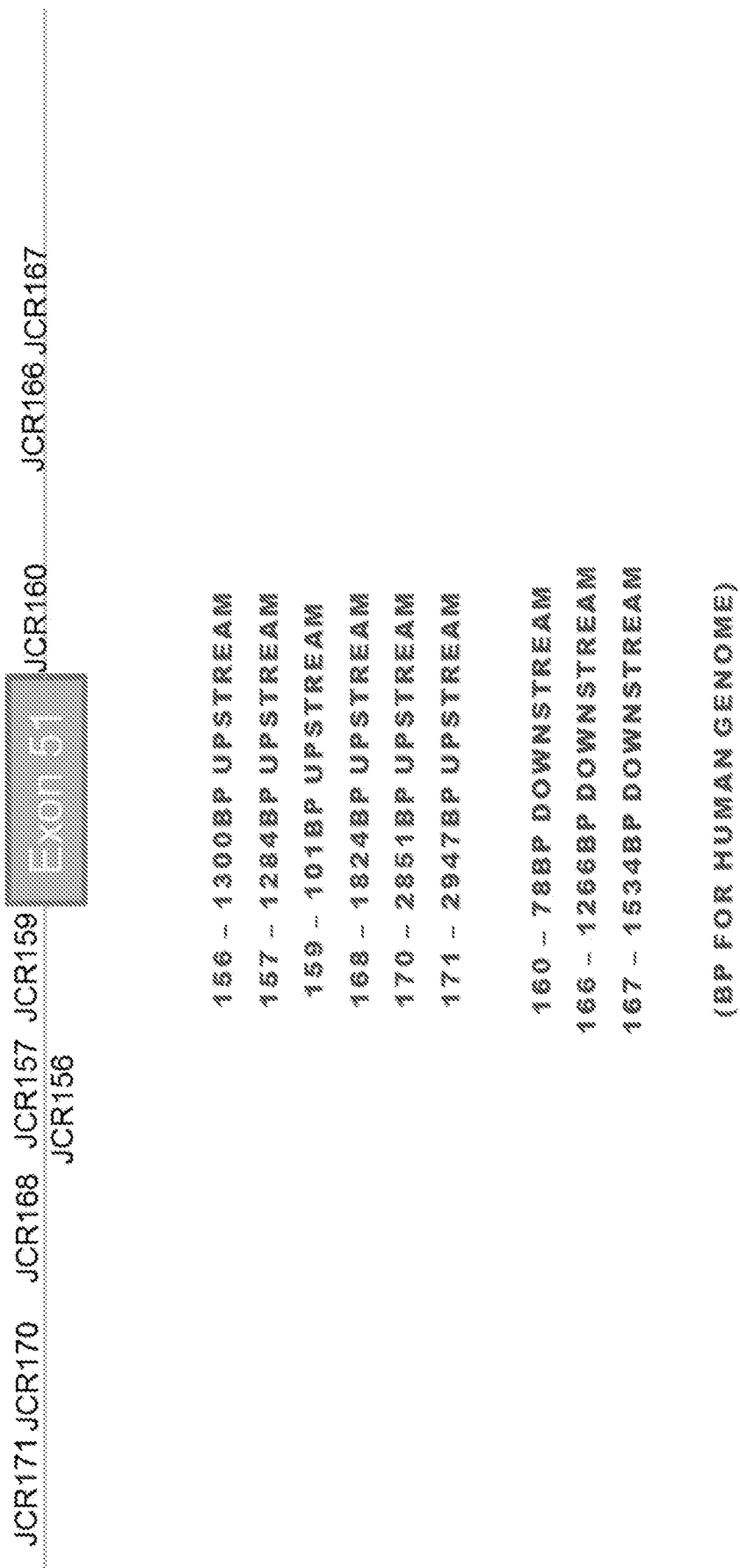


FIG. 6

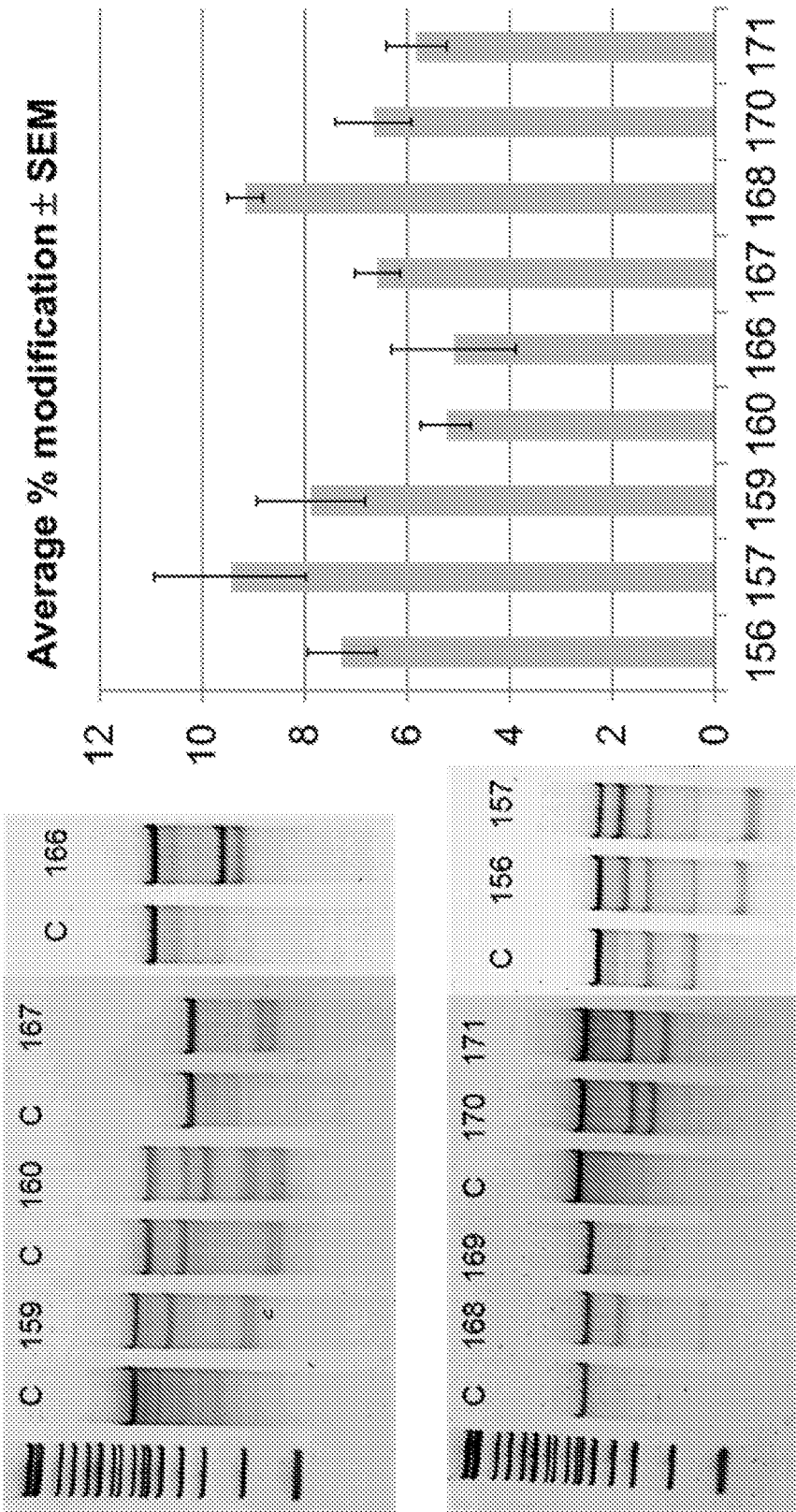


FIG. 7

| gRNA | Where | How far (nt) | % mod + SEM | Human off target (mismatches: hits in genome) | Rhesus off target |
|--------|------------|--------------|--------------|---|---|
| JCR156 | upstream | 1300 | 7.28% ± 0.66 | 4 mismatch: 23 3 mismatch: 6 2, 1 mismatch: 0 | 4 mismatch: 19 3 mismatch: 7 2 mismatch: 1 <2: 0 |
| JCR157 | upstream | 1284 | 9.45% ± 1.48 | *did not find protospacer in human genome* 4 mismatch: 6 <4 mismatch: 0 | 4 mismatch: 5 <4 mismatch: 0 |
| JCR159 | upstream | 101 | 7.88% ± 1.07 | 4 mismatch: 16 <4 mismatch: 0 | 4 mismatch: 15 3 mismatch: 1 <3 mismatch: 0 |
| JCR168 | upstream | 1824 | 9.17% ± 0.35 | 4 mismatch: 21 3 mismatch: 1 <3 mismatch: 0 | 4 mismatch: 13 <4 mismatch: 0 |
| JCR170 | upstream | 2851 | 6.66% ± 0.74 | 4 mismatch: 1 3 mismatch: 1 <3 mismatch: 0 | 4 mismatch: 3 <4 mismatch: 0 |
| JCR171 | upstream | 2947 | 5.82% ± 0.59 | 4 mismatch: 5 3 mismatch: 1 <3 mismatch: 0 | 4 mismatch: 2 3 mismatch: 1 2 mismatch: 1 <2 mismatch: 0 |
| JCR160 | downstream | 78 | 5.25% ± 0.50 | 4 mismatch: 2 <4 mismatch: 0 | 4 mismatch: 4 <4 mismatch: 0 |
| JCR166 | downstream | 1266 | 5.09% ± 1.21 | 4 mismatch: 25 3 mismatch: 5 <3: 0 | 4 mismatch: 17 3 mismatch: 5 <3 mismatch: 0 |
| JCR167 | downstream | 1534 | 6.59% ± 0.44 | 4 mismatch: 4 3 mismatch: 2 <3 mismatch: 0 | 4 mismatch: 10 <4 mismatch: 0 |

FIG. 8

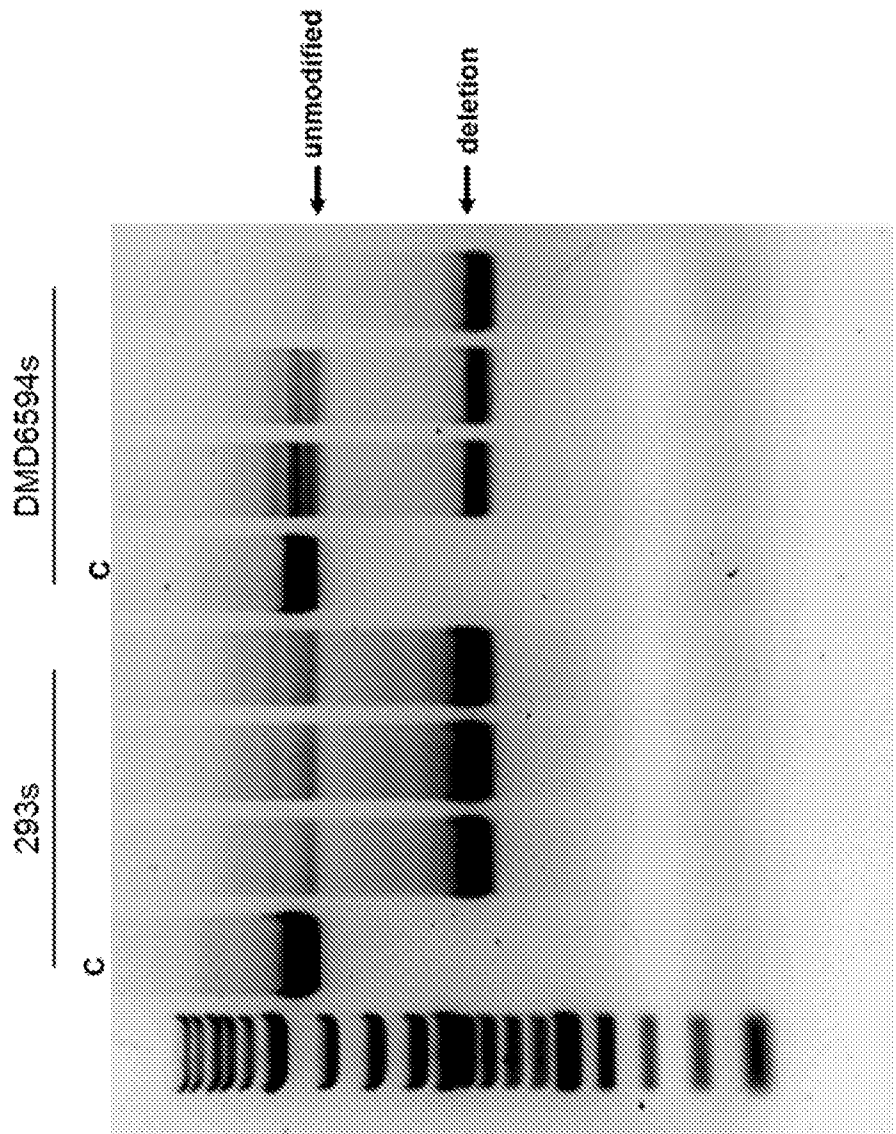


FIG. 9

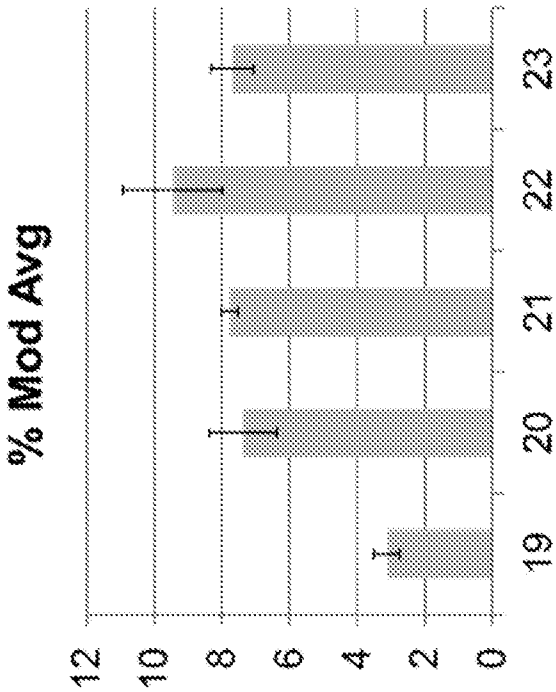


FIG. 10

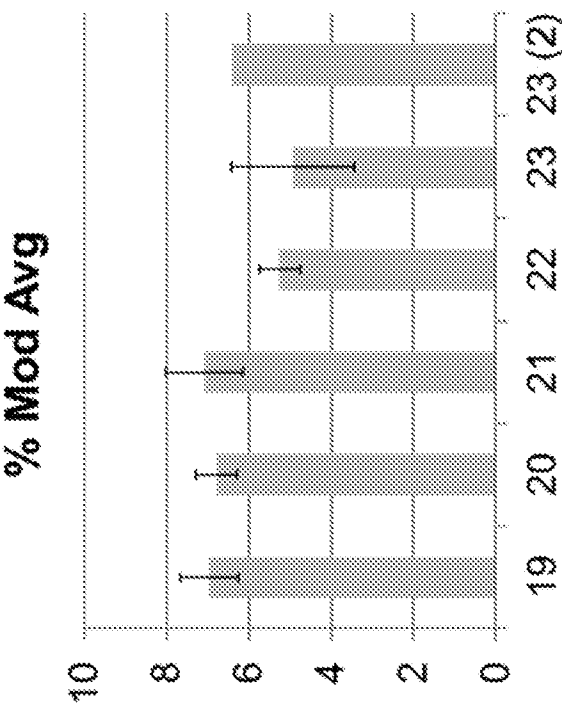
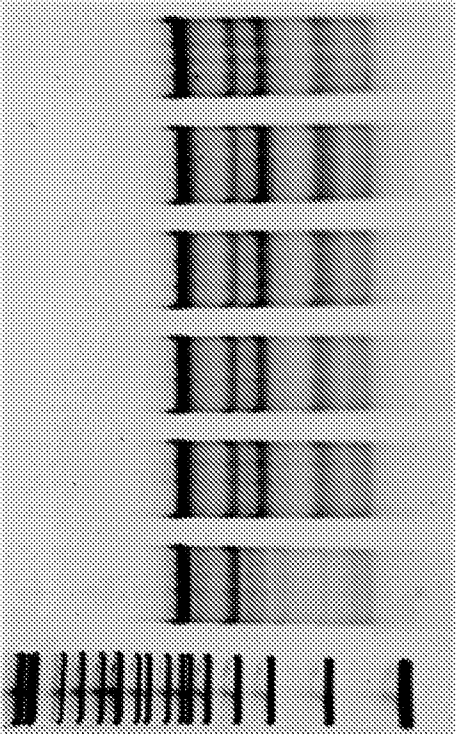
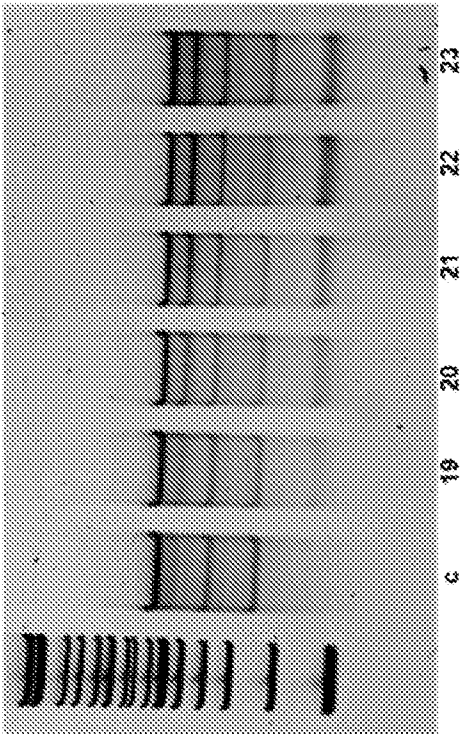


FIG. 11



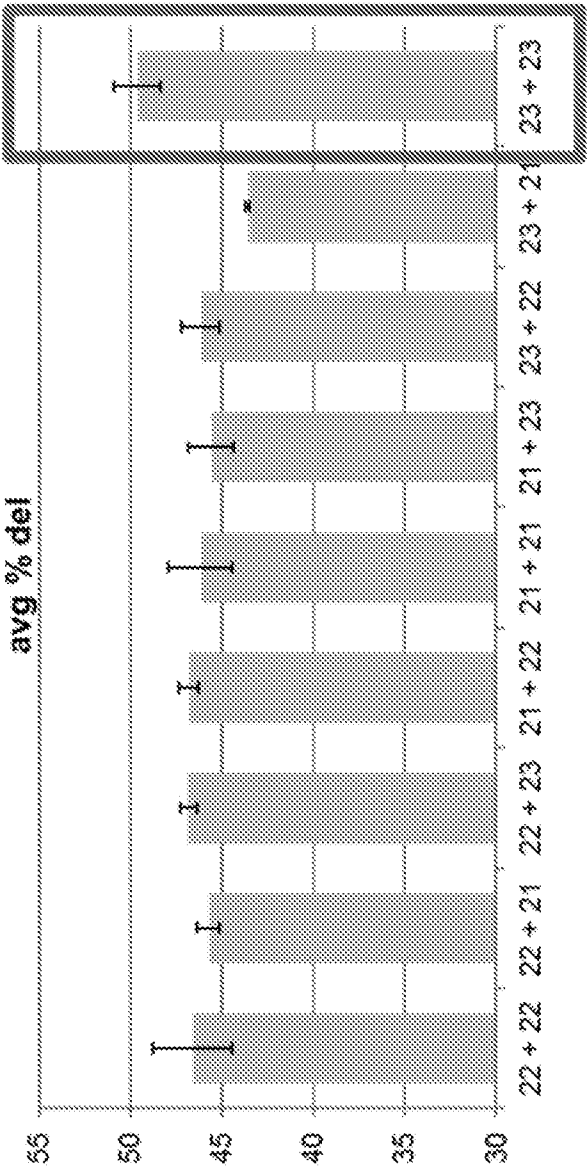
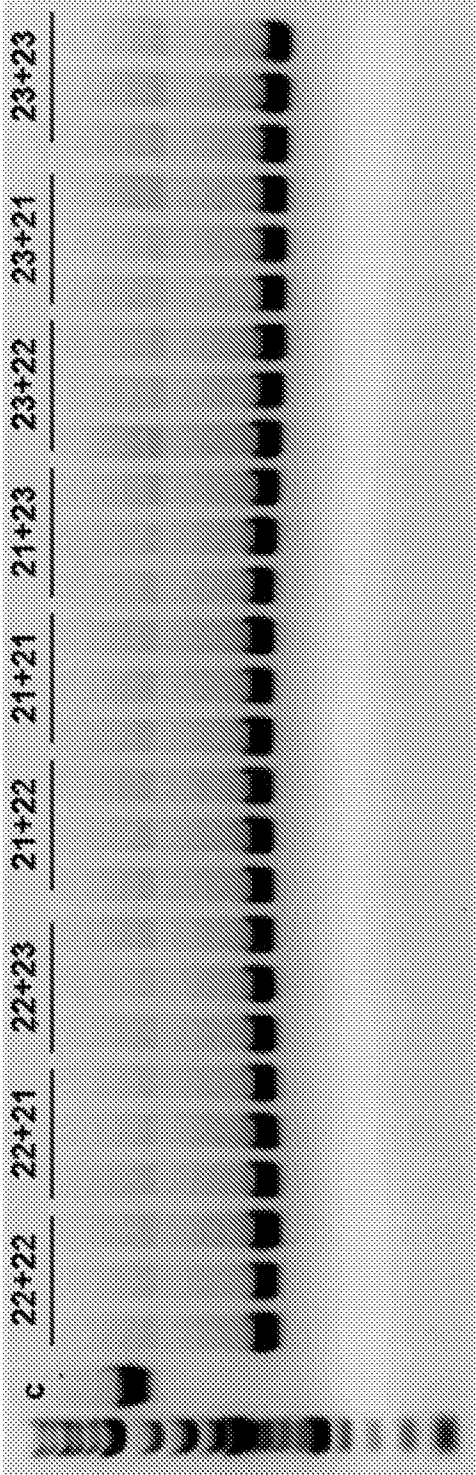


FIG. 12

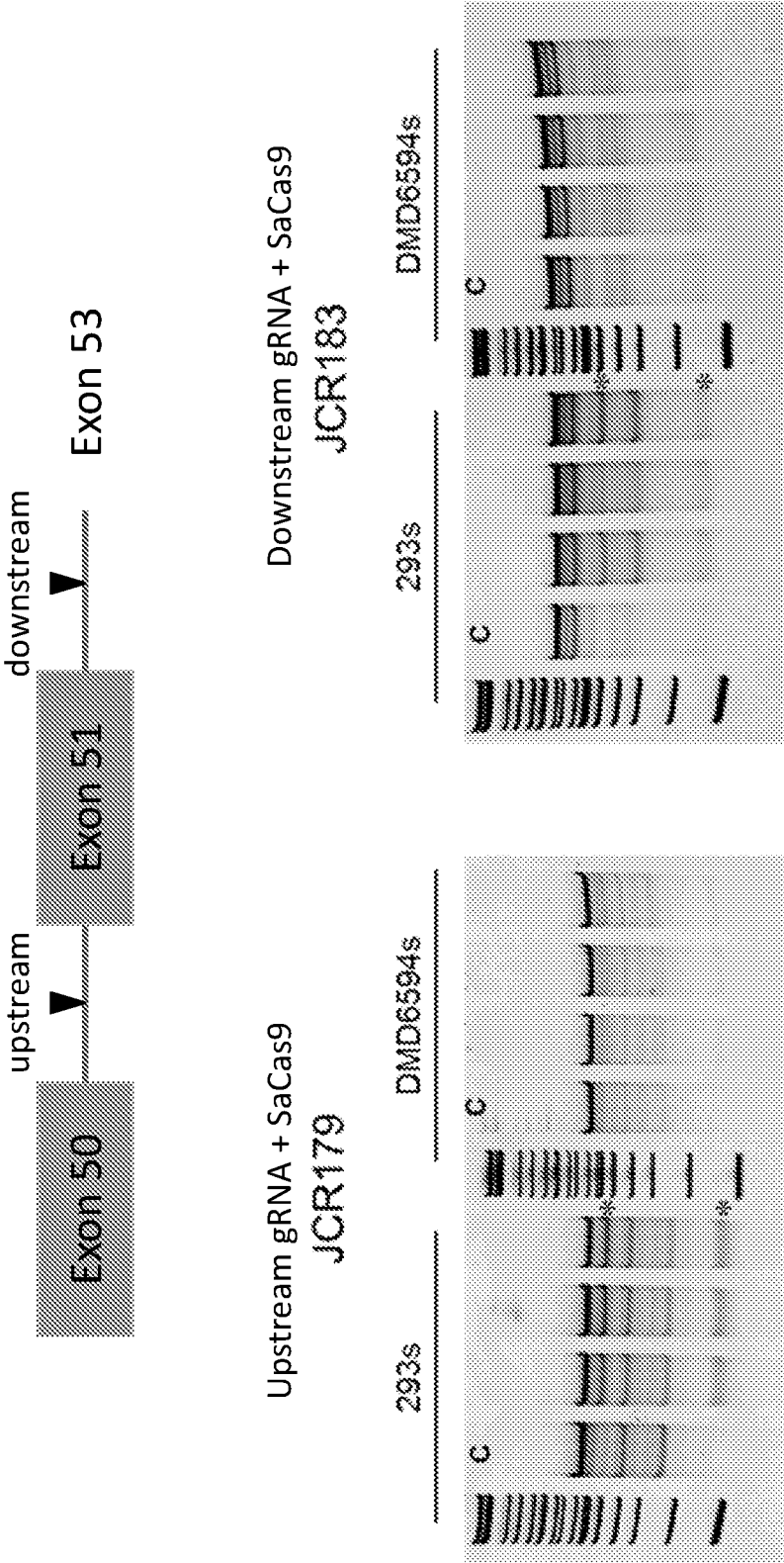


FIG. 13

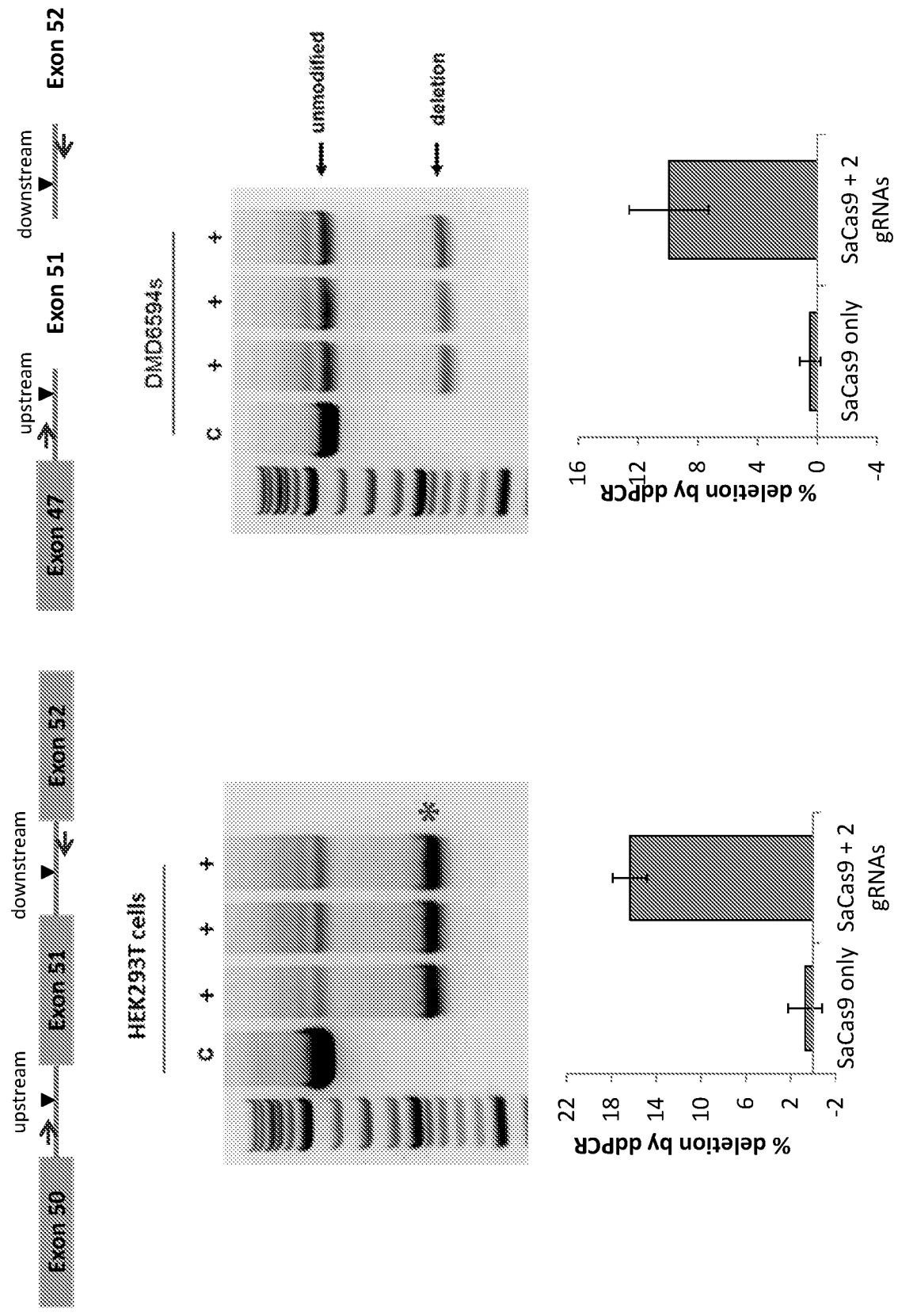


FIG. 14

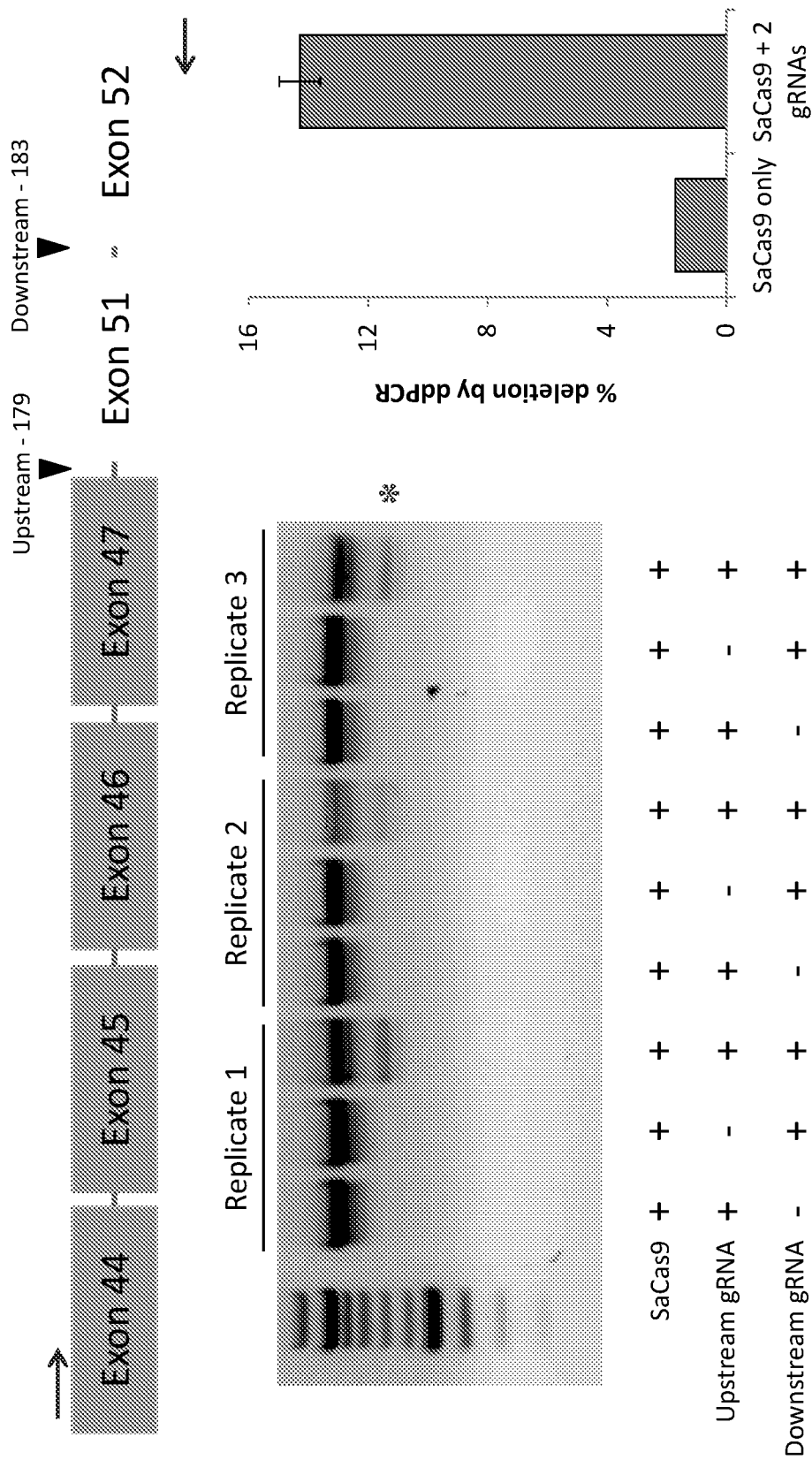


FIG. 15

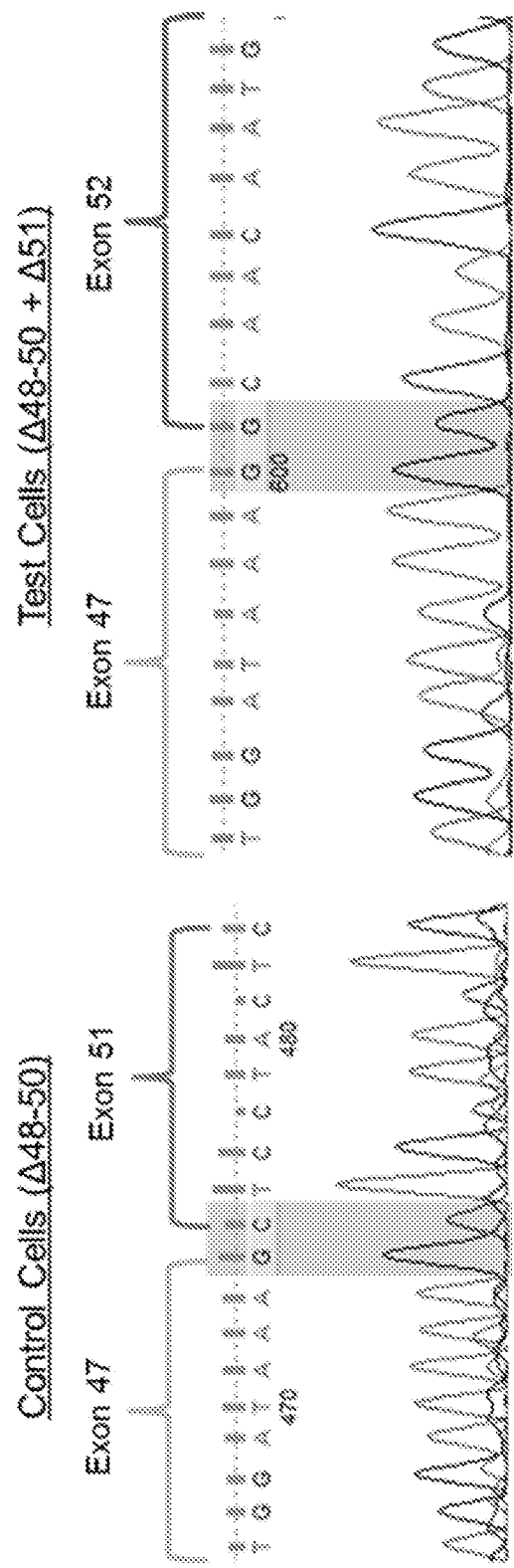


FIG. 16

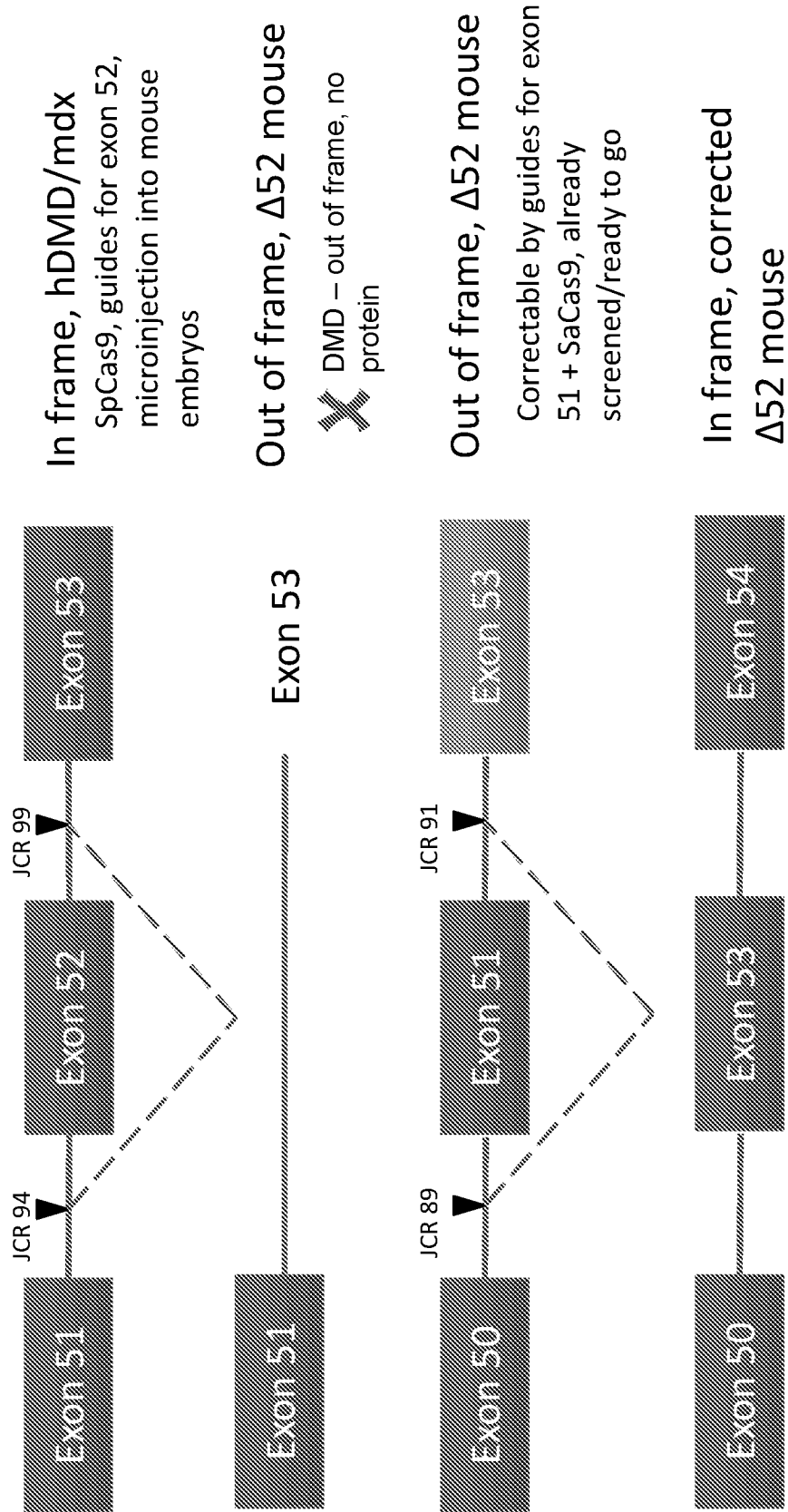


FIG. 17

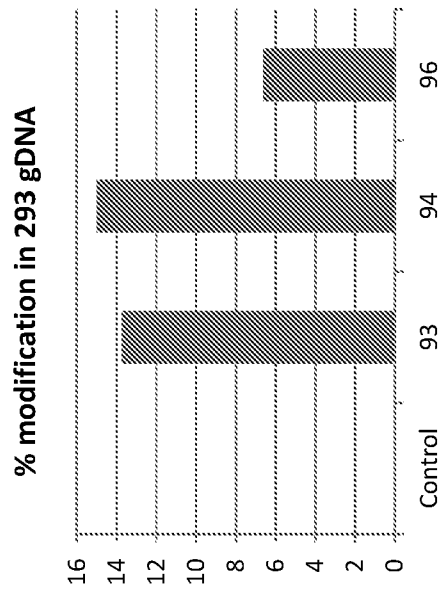
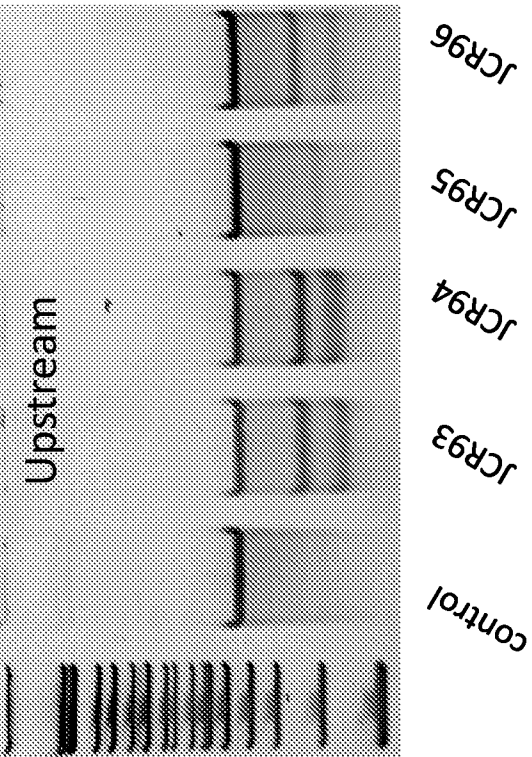
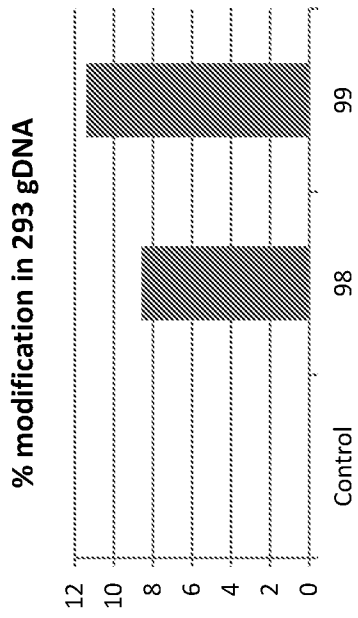
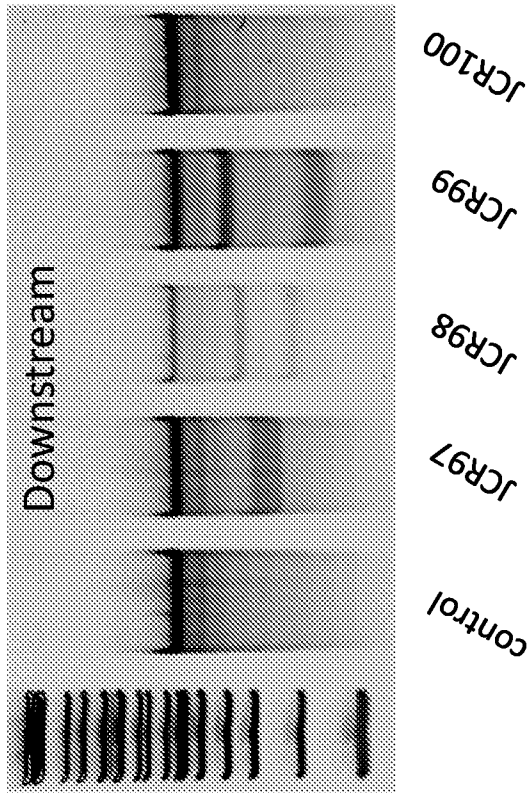


FIG. 18

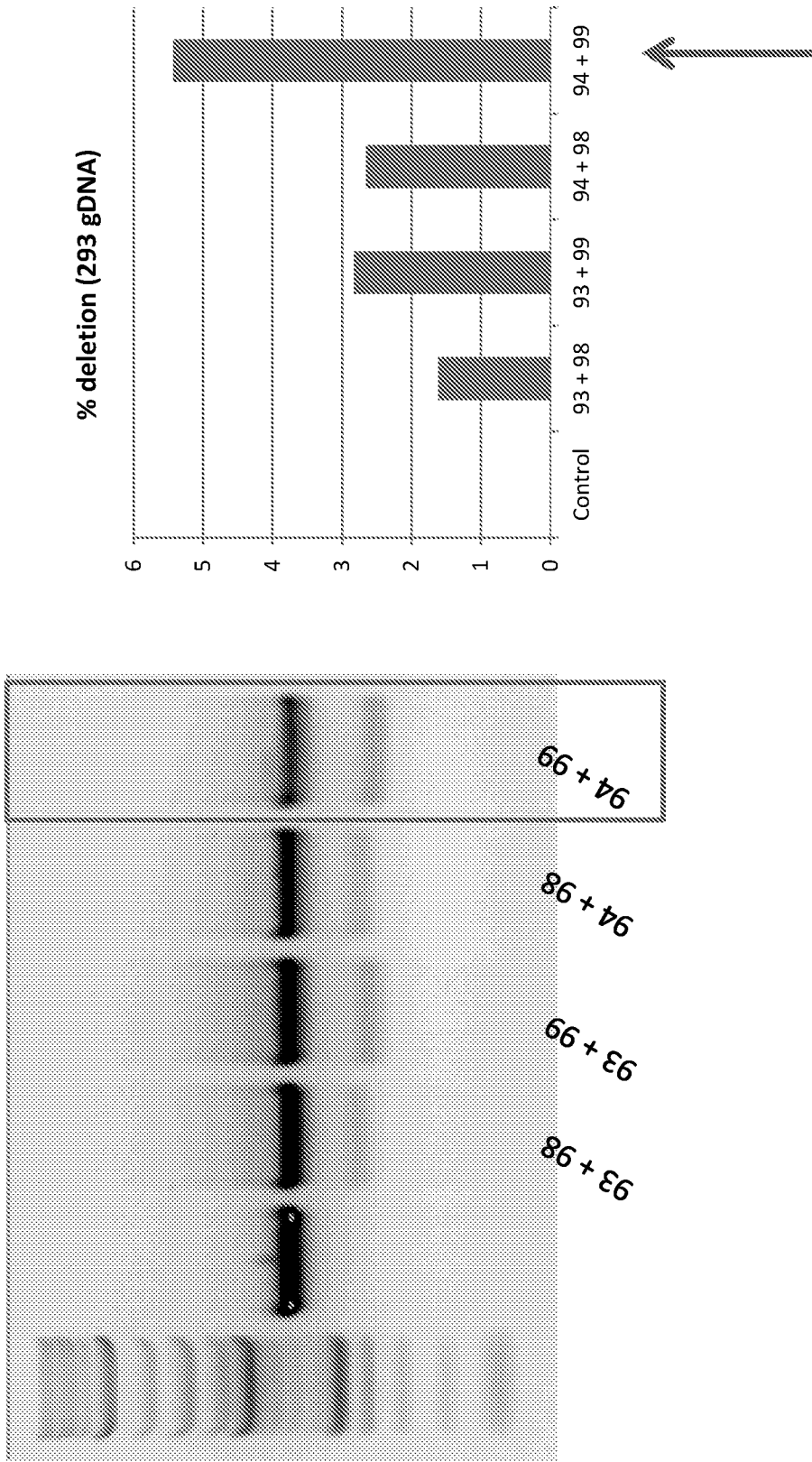


FIG. 19

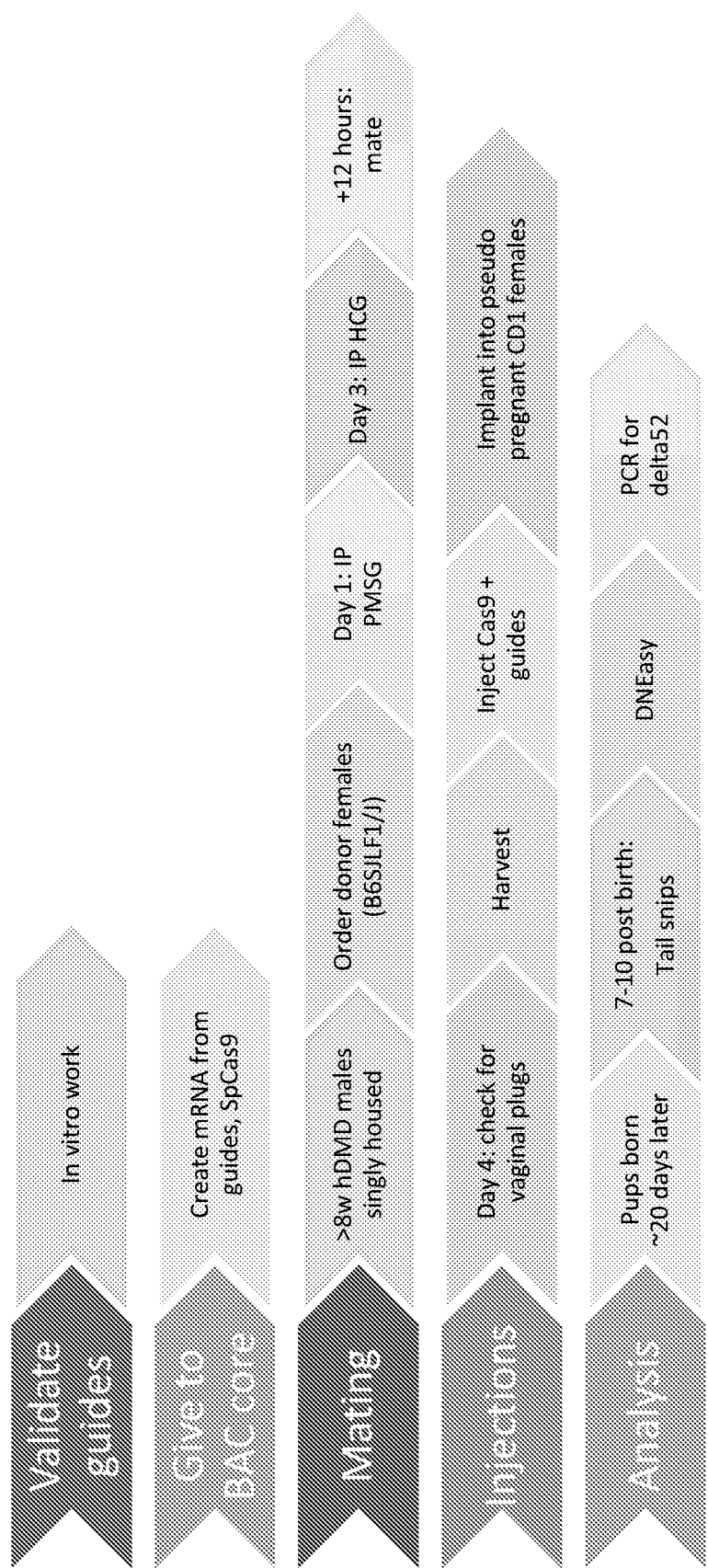


FIG. 20

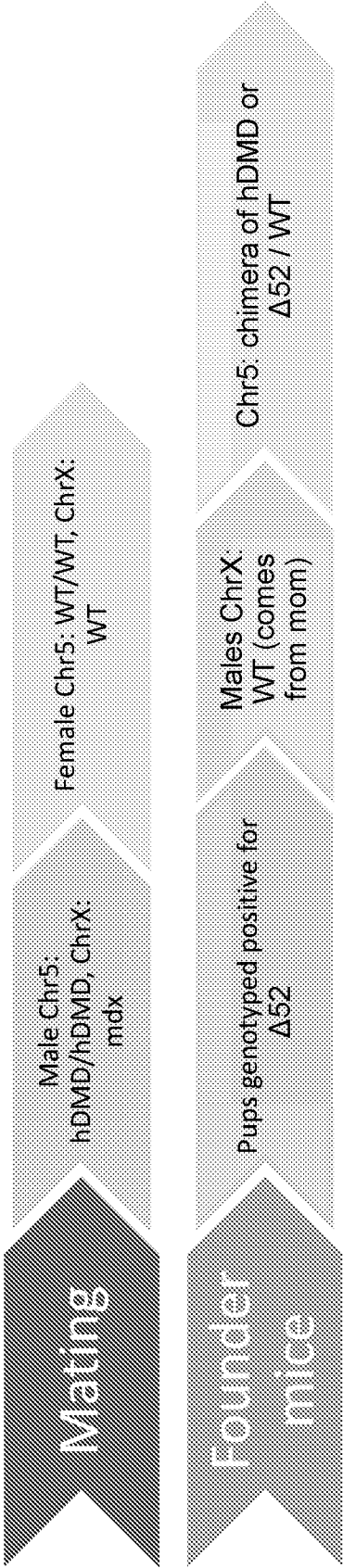


FIG. 21

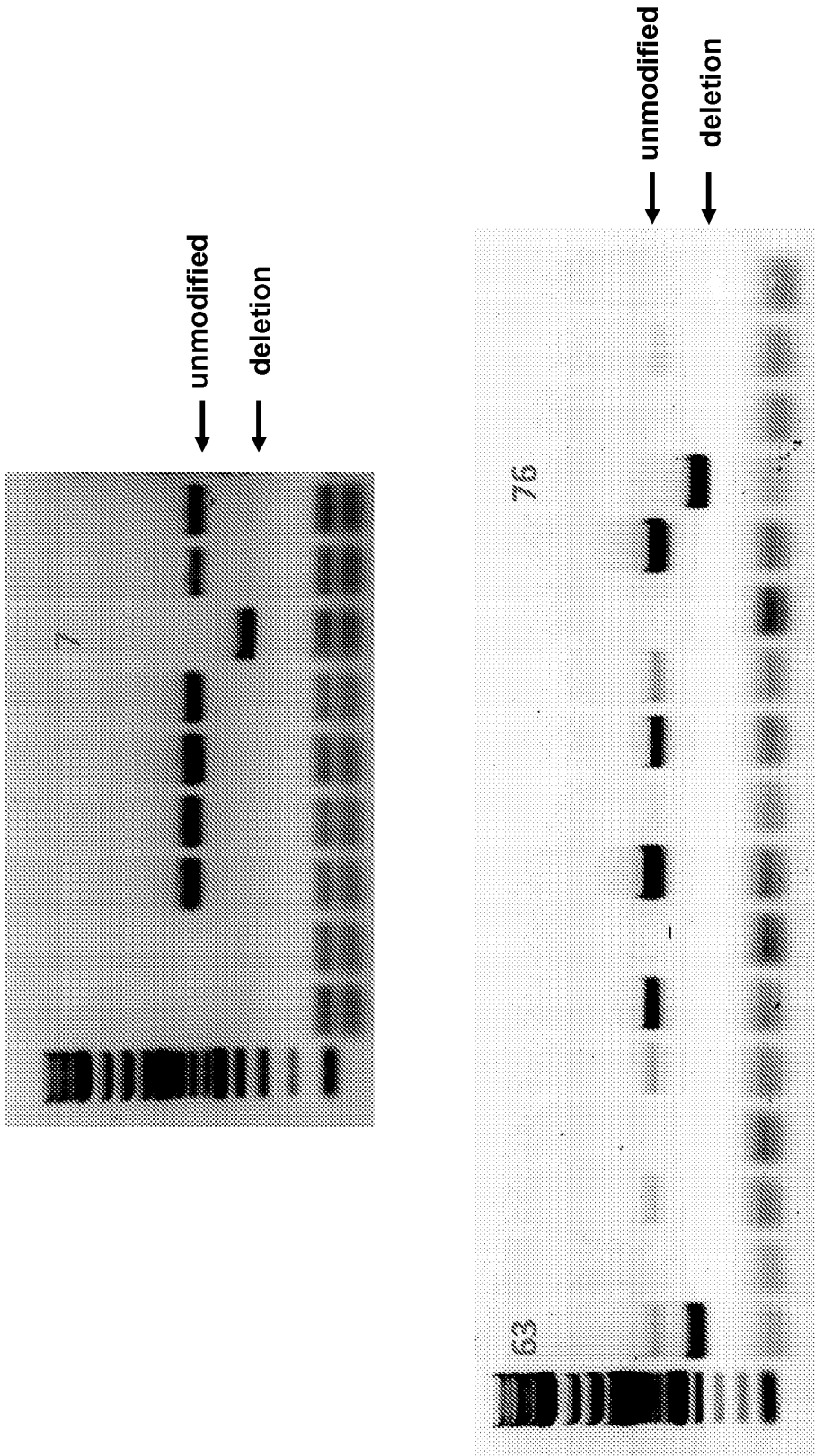


FIG. 22

Purple Bold: JCR94; **Green Underlined:** JCR99; **Grey Italicized:** Insertions/deletions

Delta 52: AACGCTGAAGAACCCCT**GAT**AATAGAGAGAAATACATTTTTAAATCAATTCAGG (SEQ ID NO: 46)

Founder 76: AAGAACCCCT**GAT**A ~ TTATCTTAGTAGATTA ~ ATAGAGAGAAATACATTTTTTAAA (SEQ ID NO: 47)

Founder 63: GCTGAAGAACCCCT**GA** ~ AAAAATACATTTTTTATCAATTCAGG (SEQ ID NO: 48)

Founder 7: GAAT ~ (19x) ~ GATTTTCTTGAGAGAAATACCAAT ~ AAATC (SEQ ID NO: 49)

FIG. 23

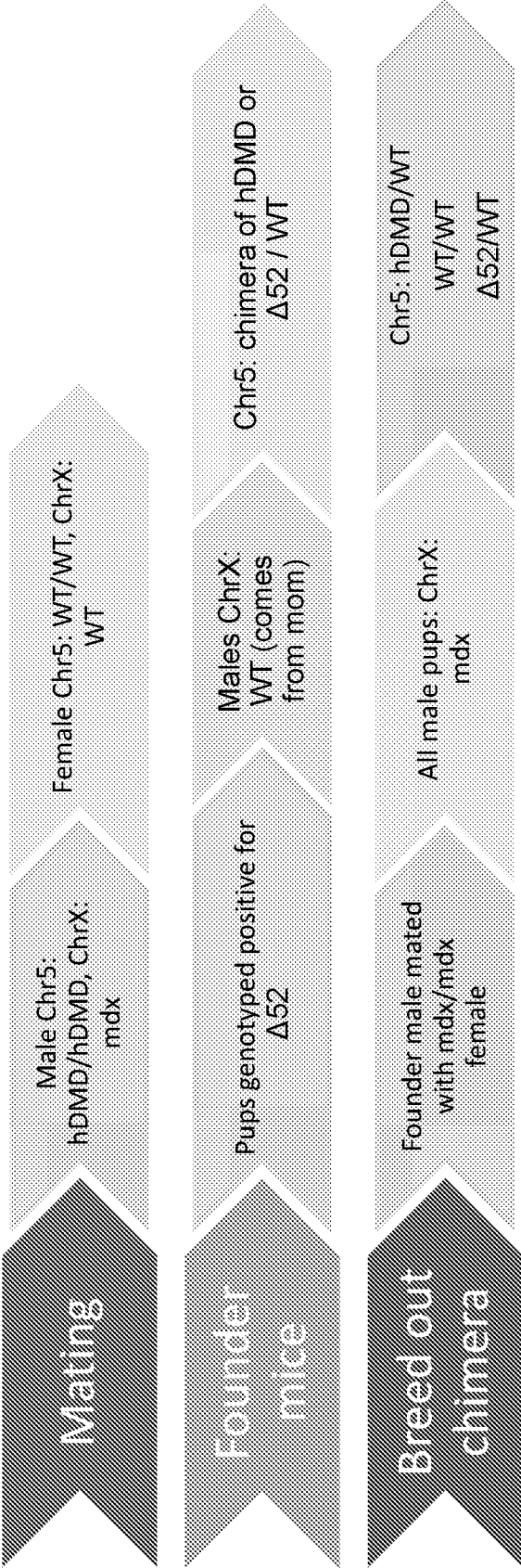


FIG. 24

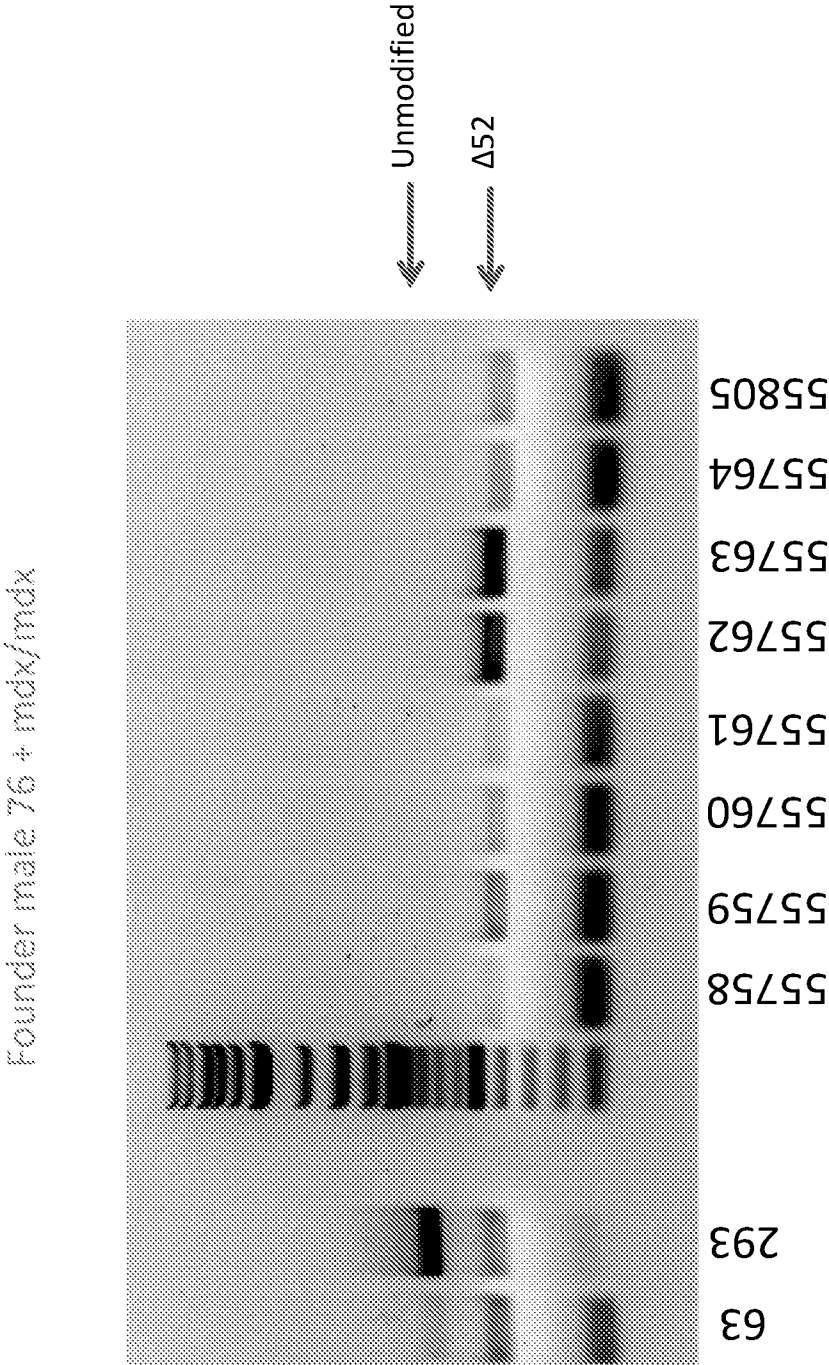
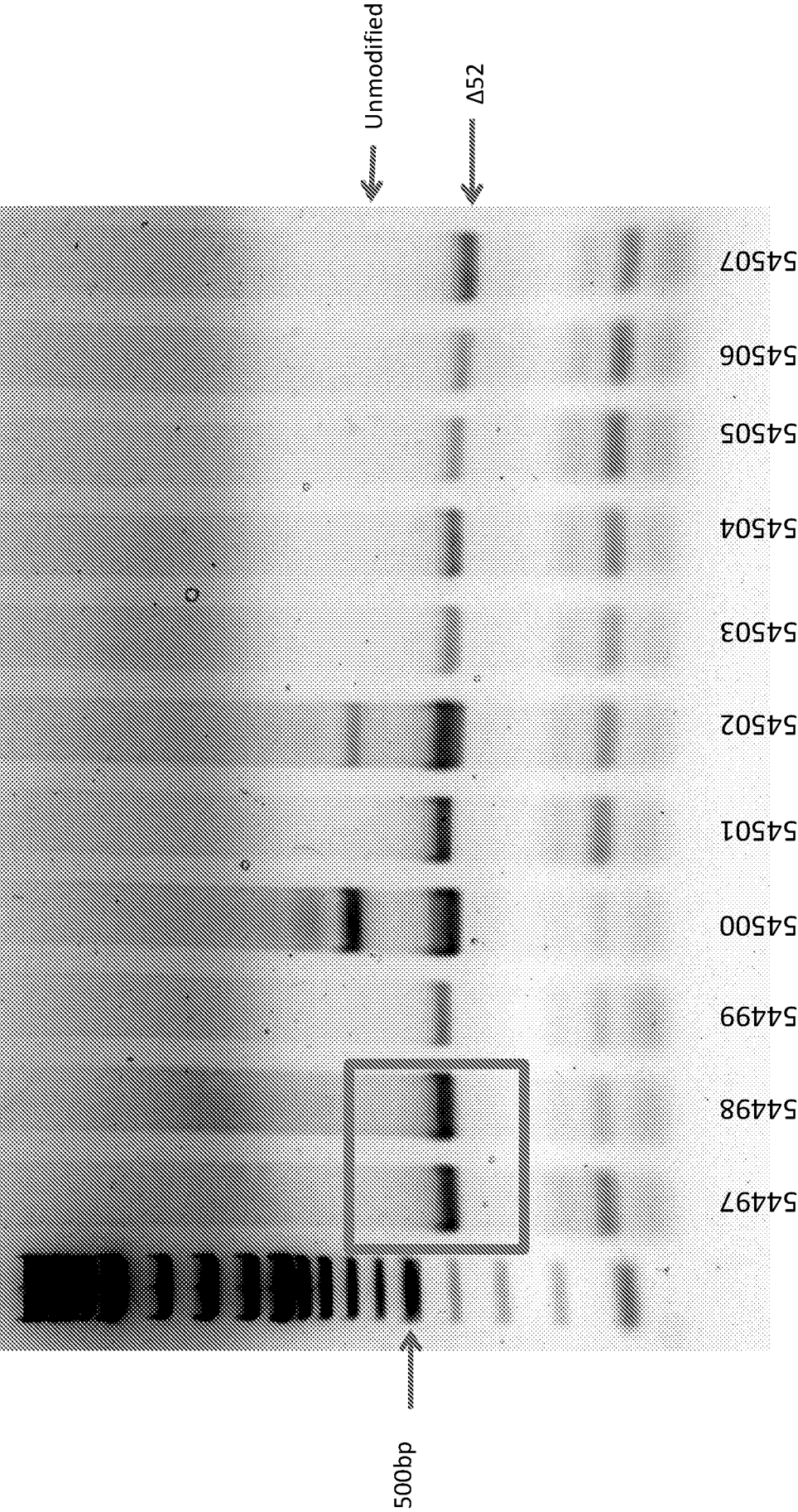


FIG. 25



Purple Bold: JCR94; Green Underlined: JCR99; Grey Italicized: Insertions/deletions

Delta 52: AACGCTGAAGAACCCT**GAT**AAATAGAAAGAAATACATTTTAAATCAATTCAGG (SEQ ID NO: 50)

54497: AACGCTGAAGAACCCT**GAT**ATTATCTTAGTAGATTAAATAGAAAGAAATACATTTTTAAAT (SEQ ID NO: 51)

54498: AACGCTGAAGAACCCT**GAT**ATTATCTTAGTAGATTAAATAGAAAGAAATACATTTTTAAAT (SEQ ID NO: 52)

FIG. 27

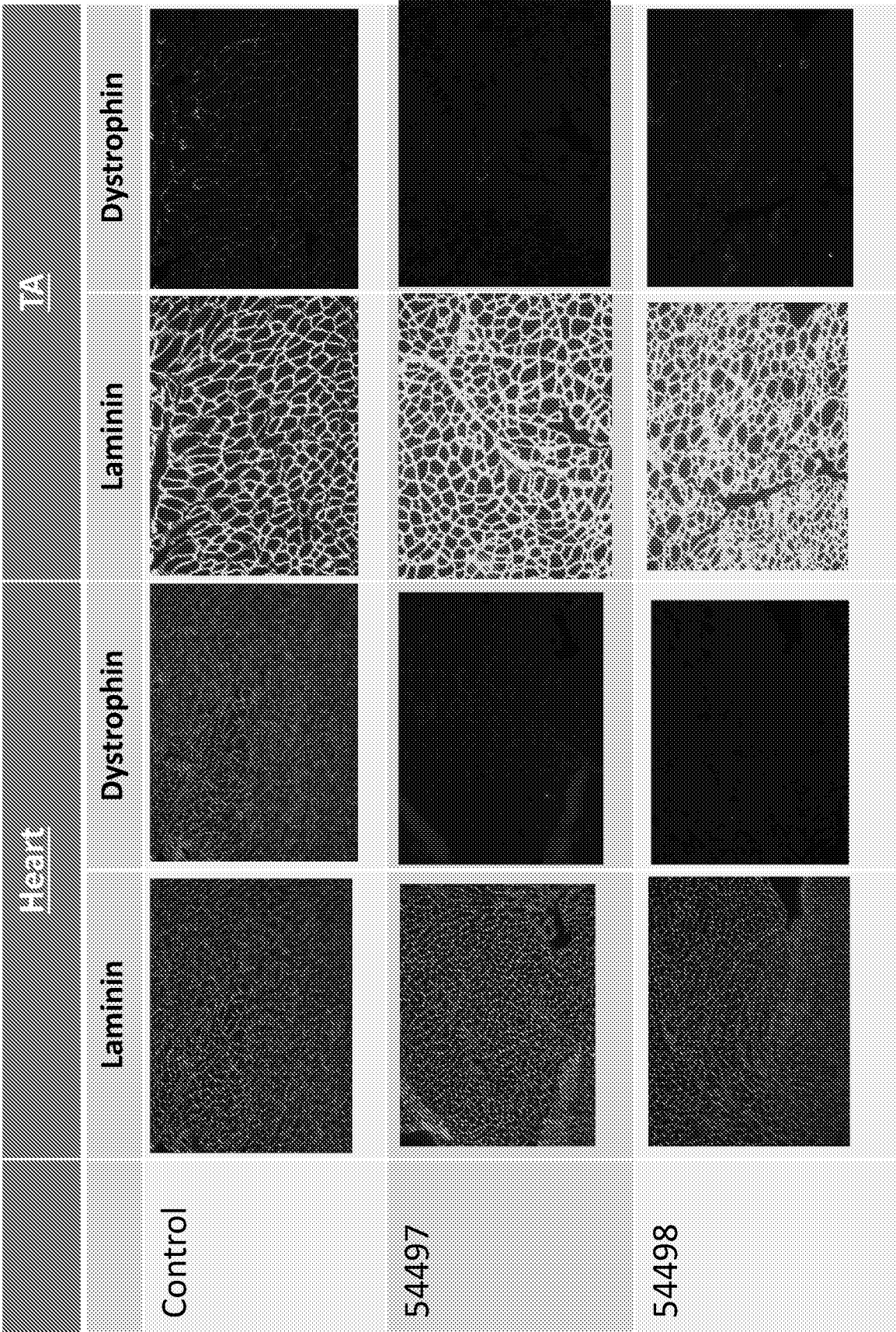


FIG. 28

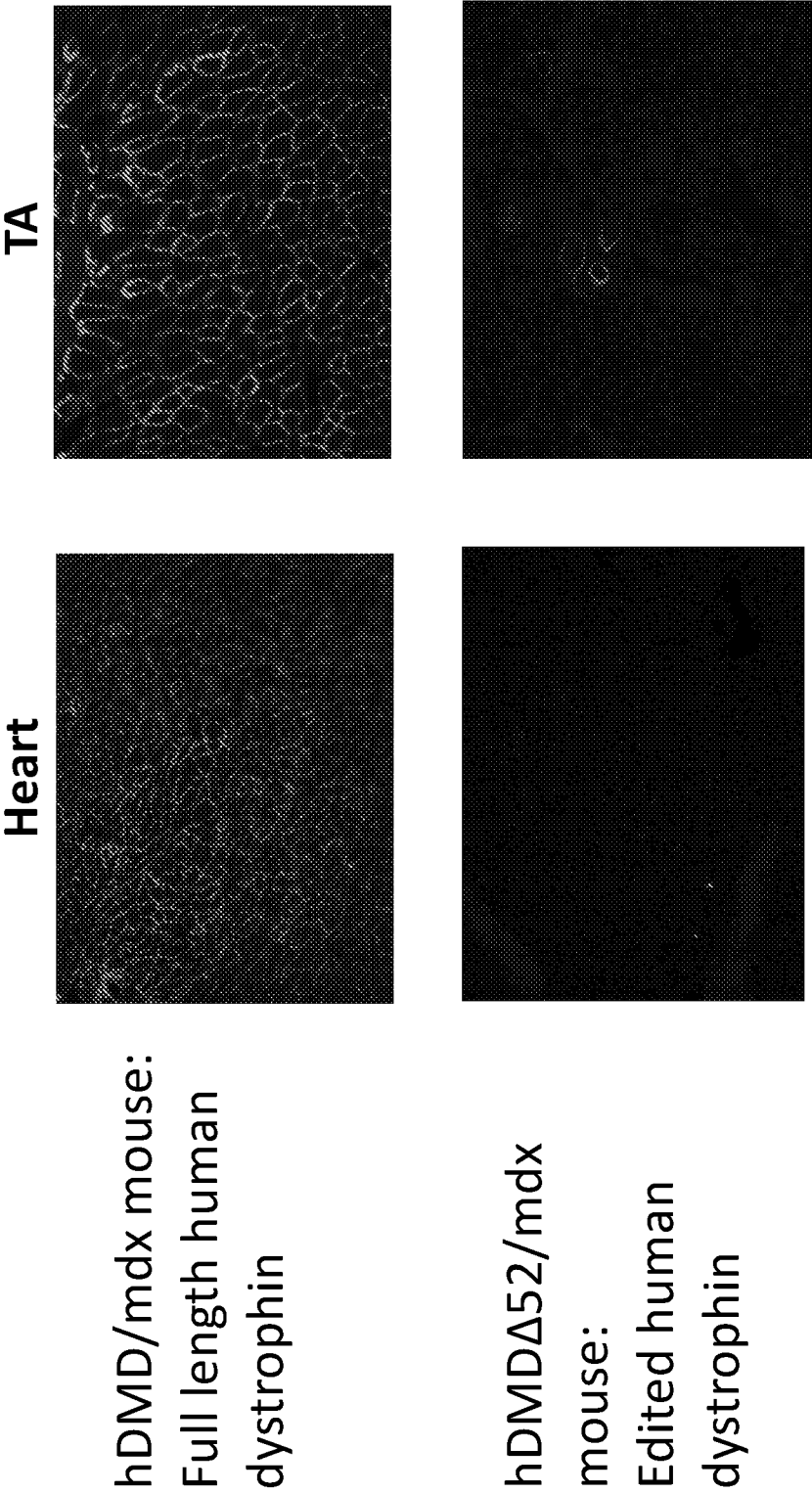


FIG. 29

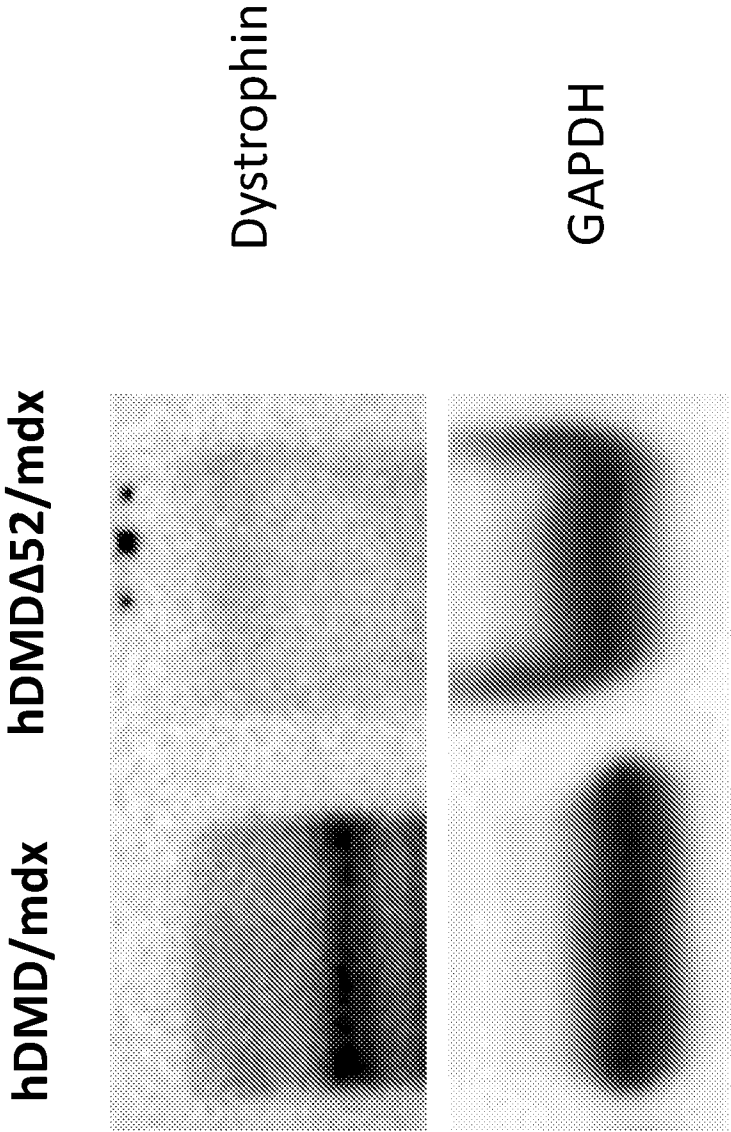
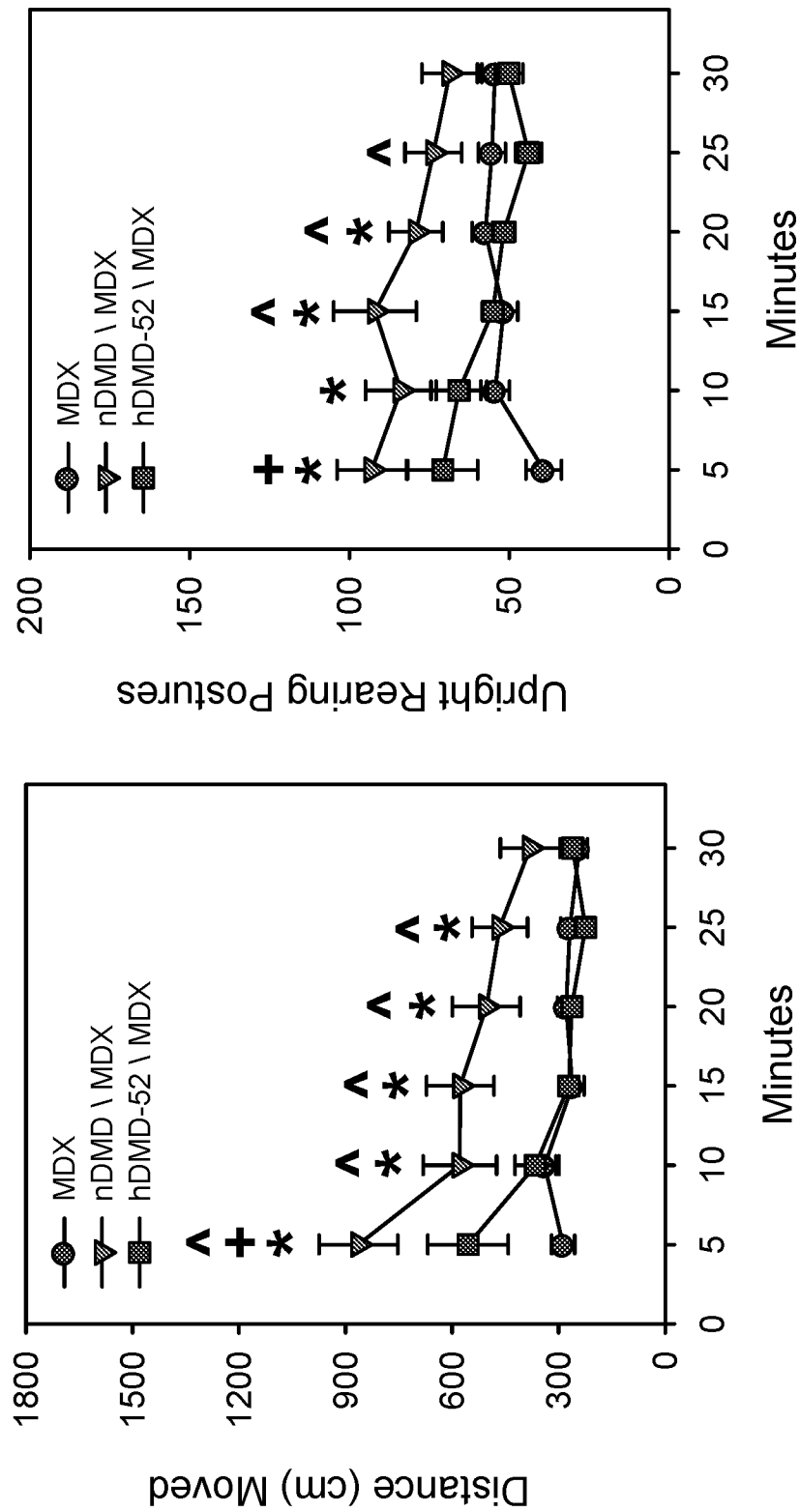


FIG. 30



^ p<0.05 Δ52/mdx mice compared to hDMD/mdx mice
* p<0.05 mdx mice compared to hDMD/mdx mice
+ p<0.05 mdx mice compared to Δ52/mdx

FIG. 31

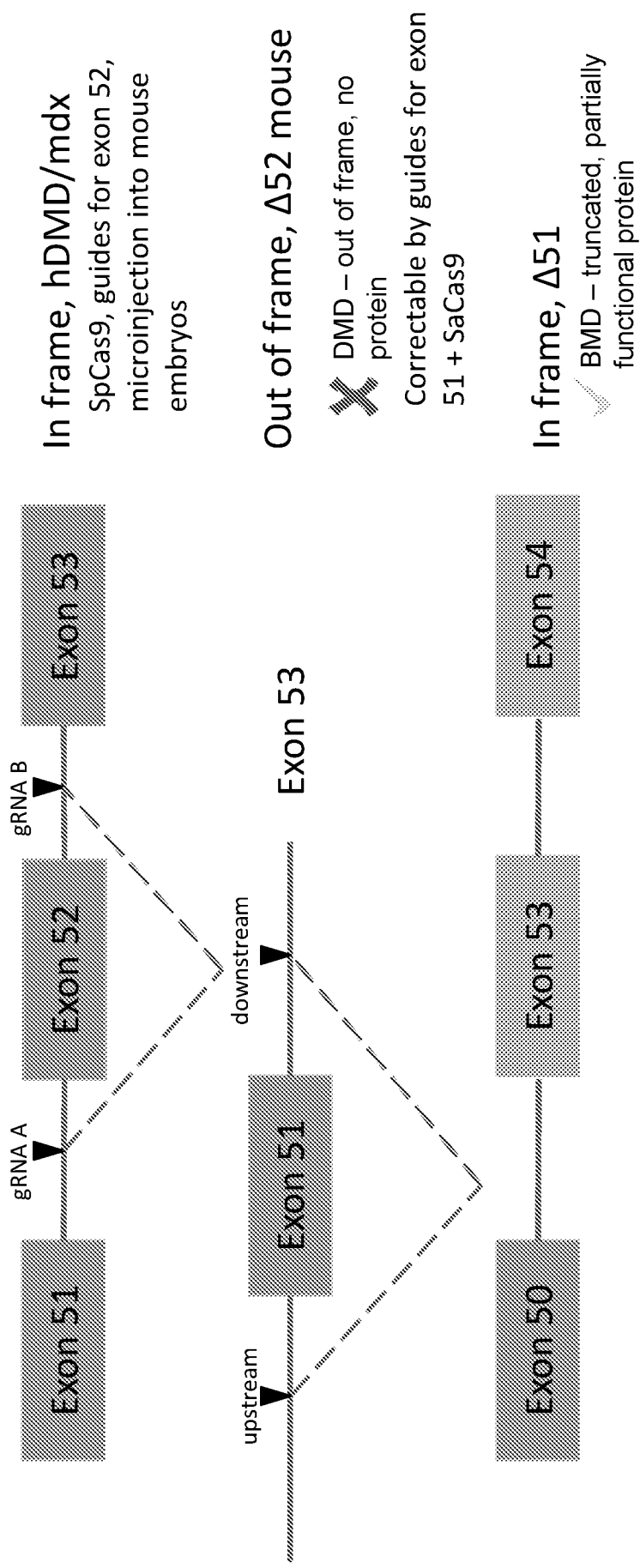
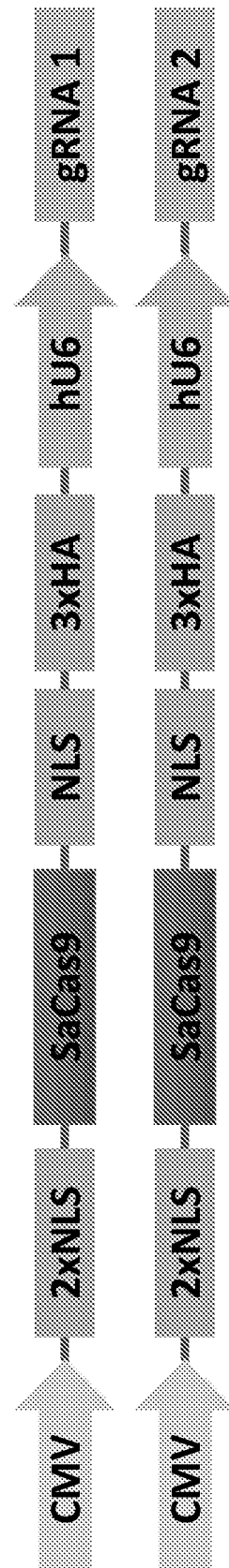
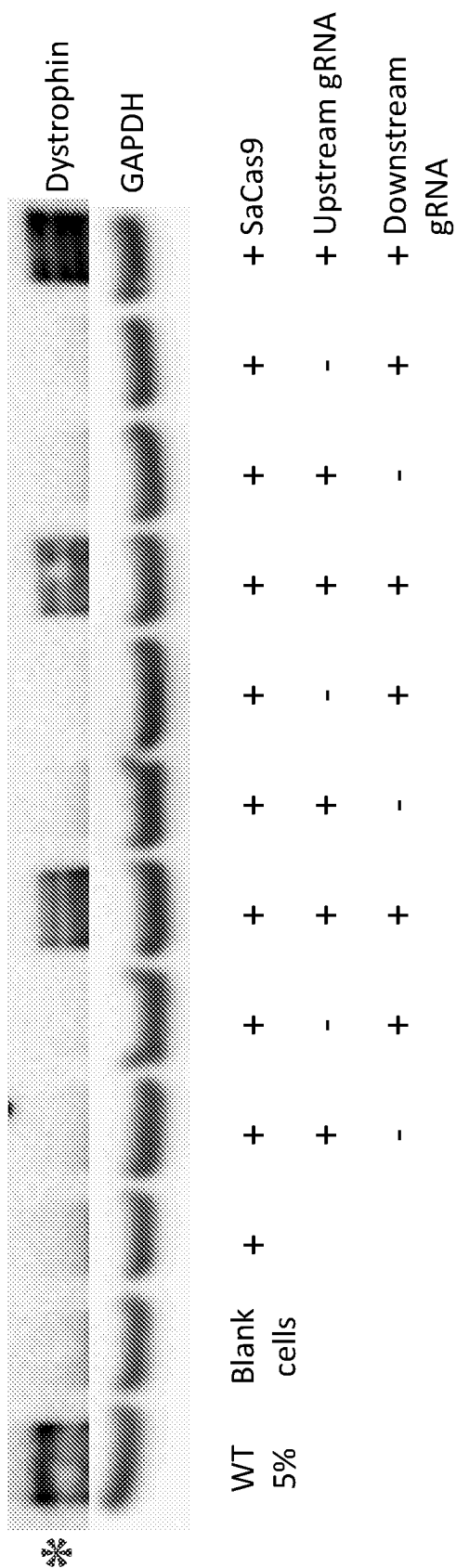


FIG. 32



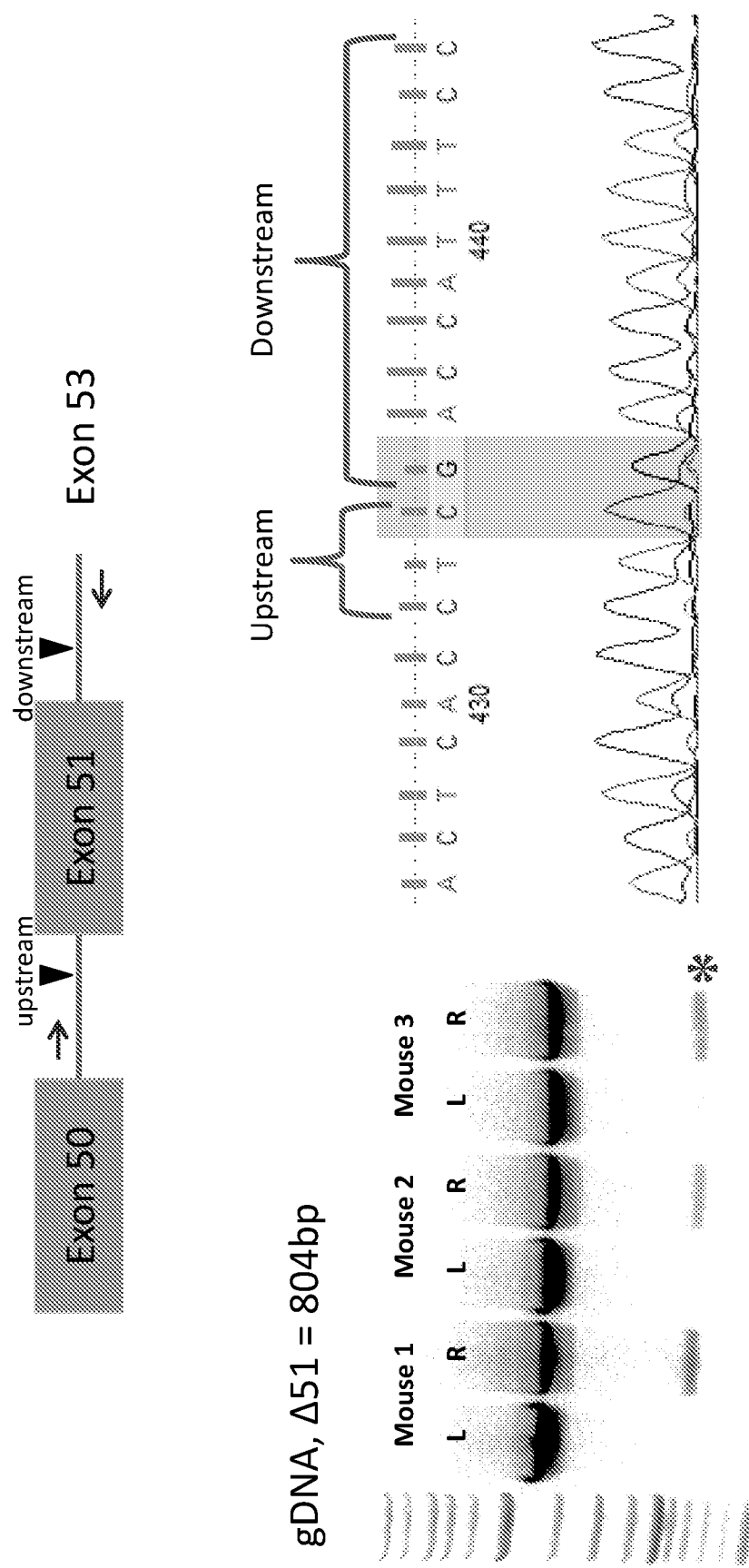
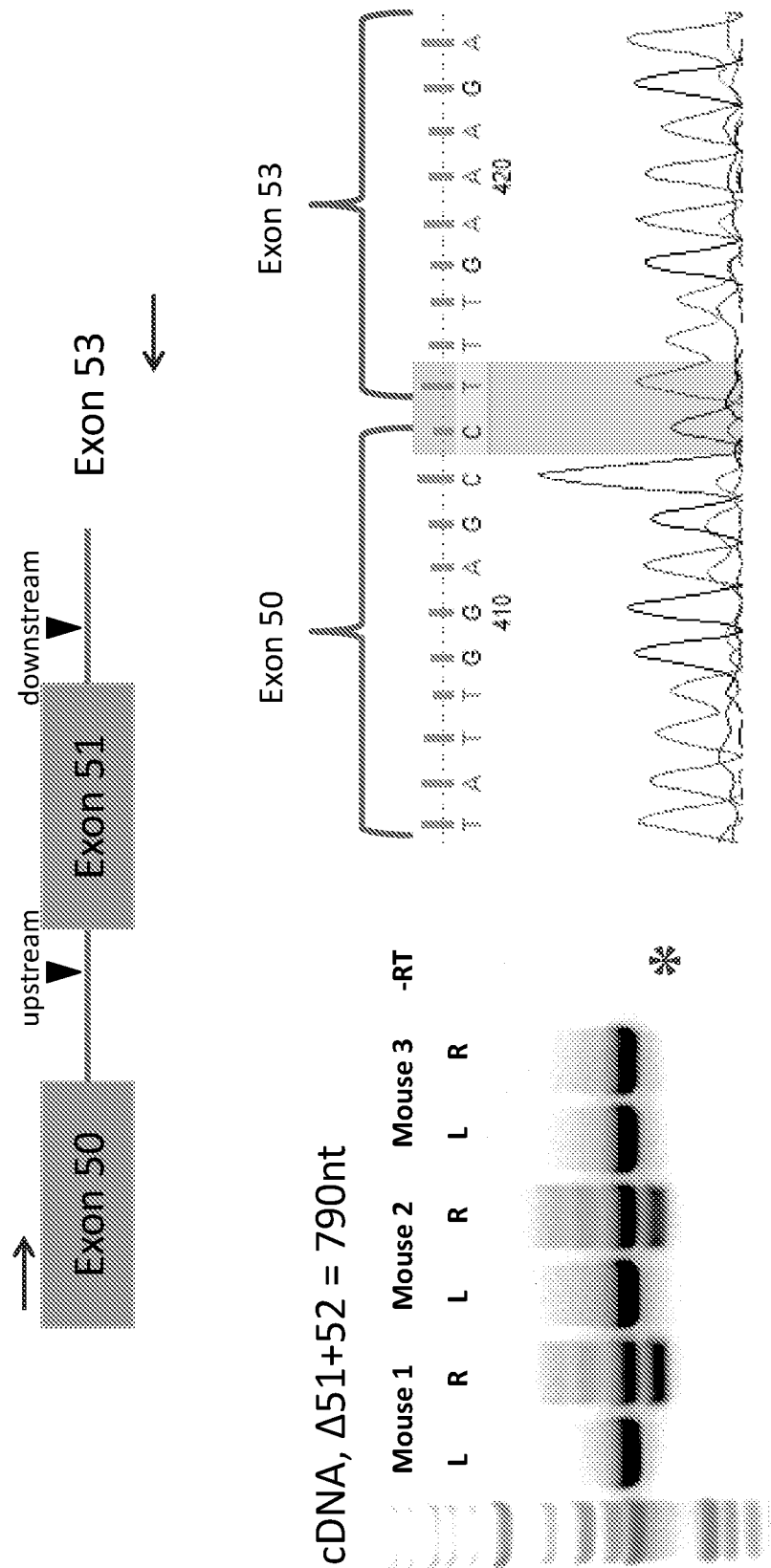


FIG. 35



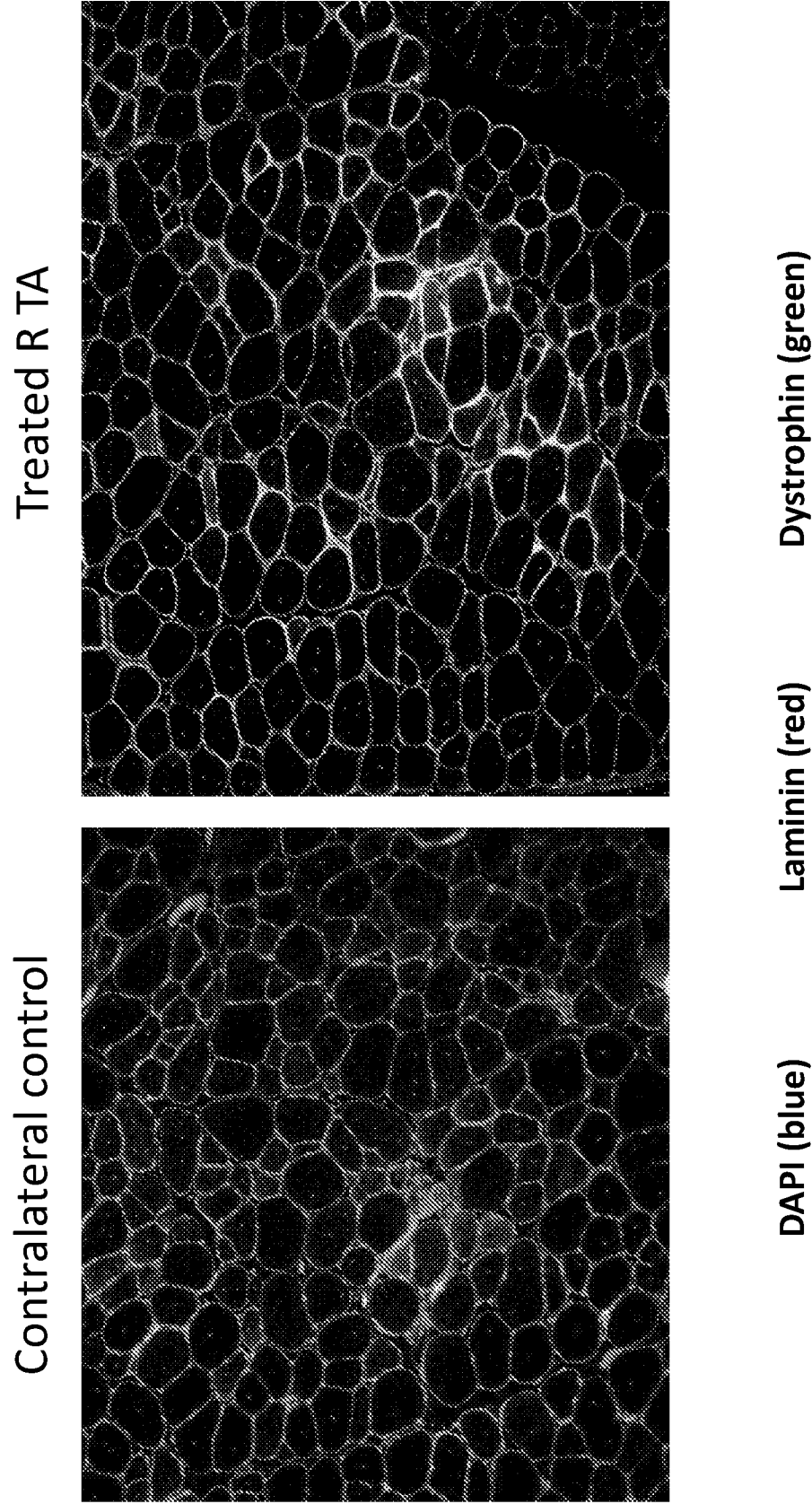


FIG. 37

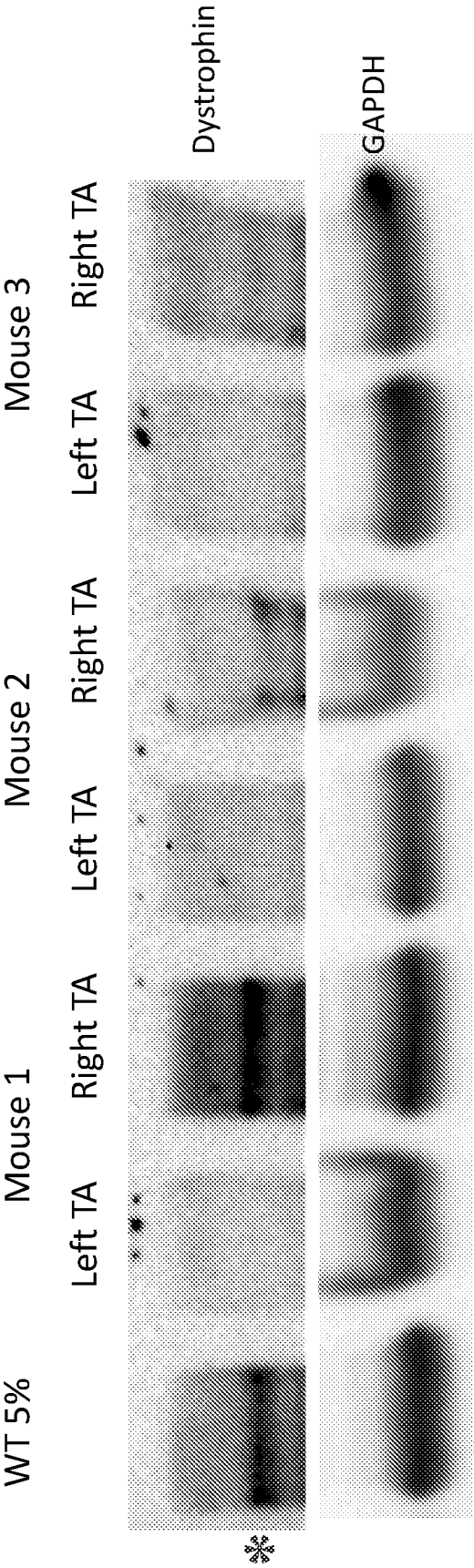


FIG. 38

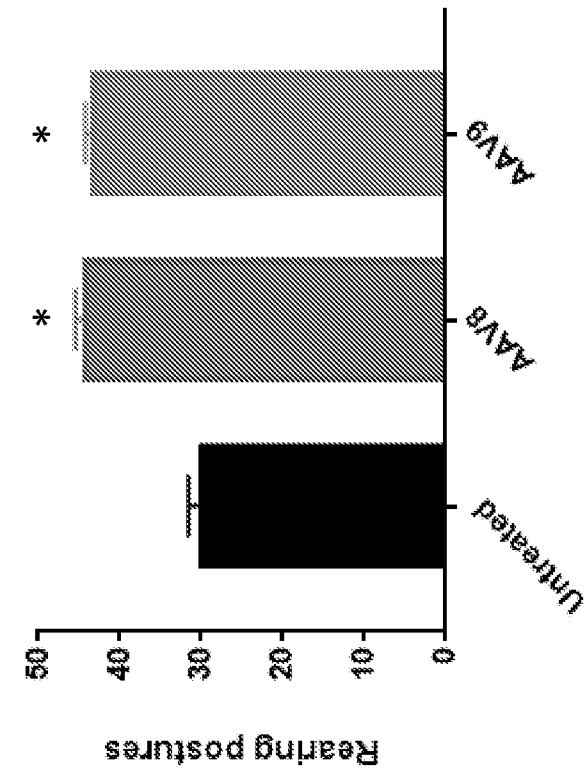


FIG. 40

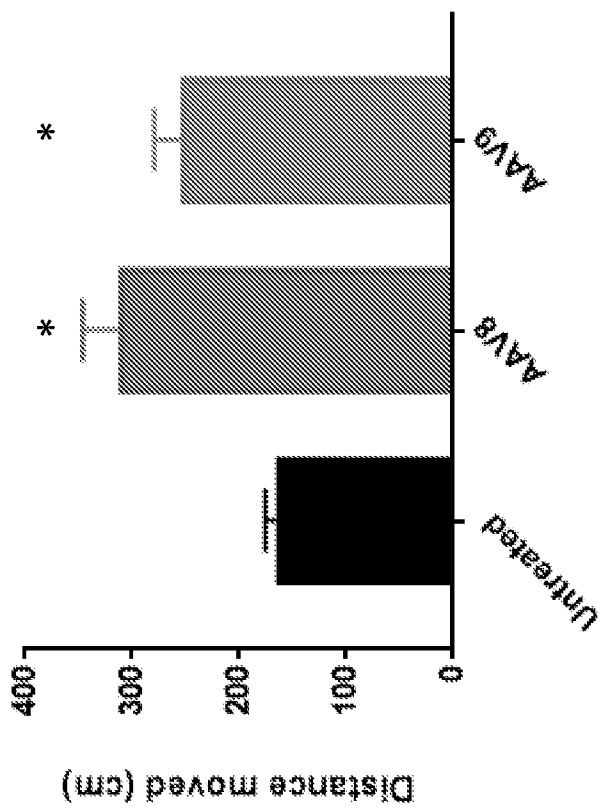


FIG. 39

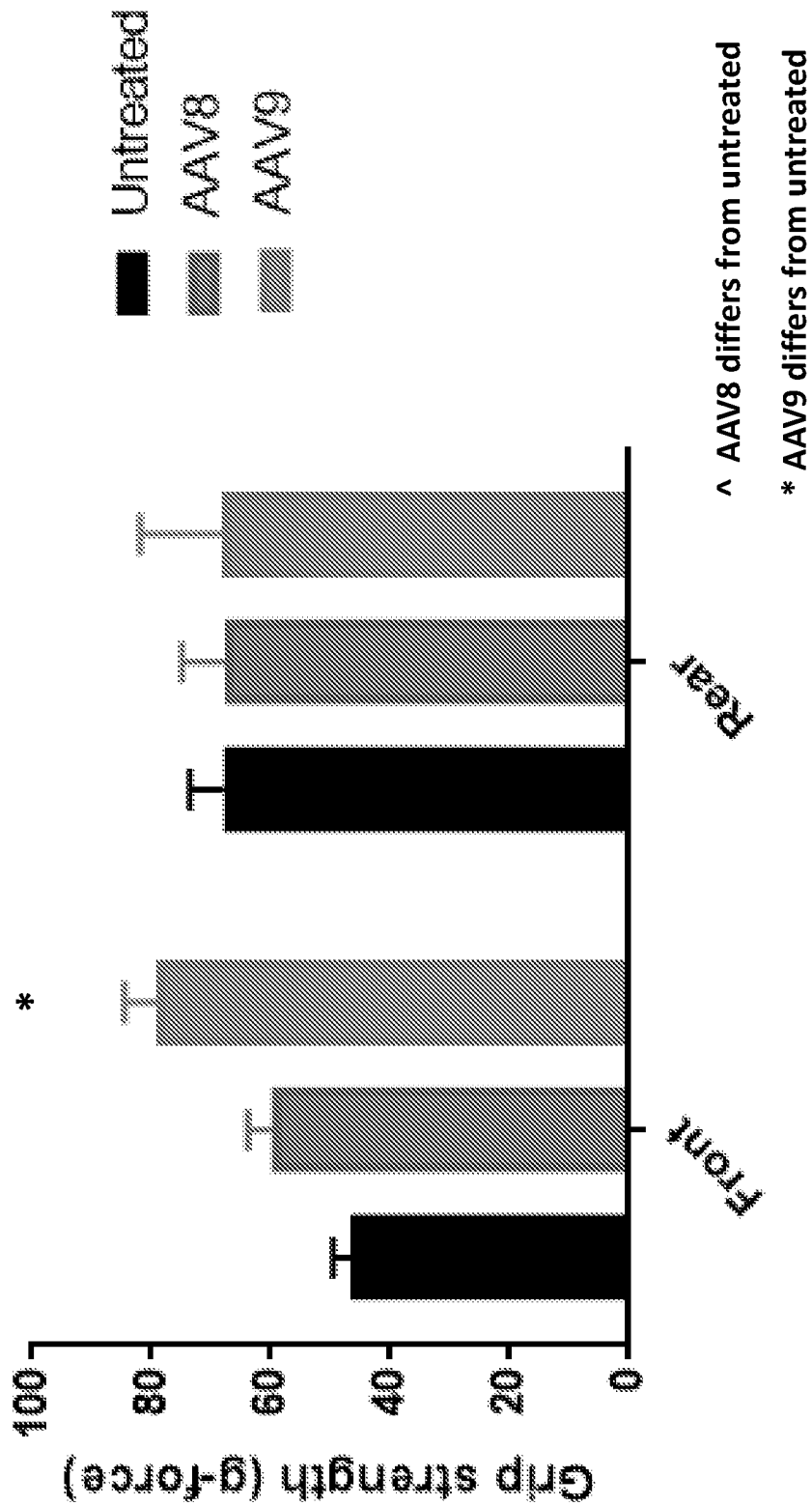


FIG. 41

PATENT COOPERATION TREATY

PCT

DECLARATION OF NON-ESTABLISHMENT OF INTERNATIONAL SEARCH REPORT

(PCT Article 17(2)(a), Rules 13^{ter}.1(c) and (d) and 39)

| | | |
|---|---|--|
| Applicant's or agent's file reference 028193-9251-WO00 | IMPORTANT DECLARATION | Date of mailing (<i>day/month/year</i>) 06 APR 2017 |
| International application No. PCT/US16/64285 | International filing date (<i>day/month/year</i>) 30 November 2016 | (Earliest) Priority Date (<i>day/month/year</i>) 30 November 2015 |
| International Patent Classification (IPC) or both national classification and IPC IPC: C12Q 1/68; CPC: C12Q 2600/158 | | |
| Applicant DUKE UNIVERSITY | | |

This International Searching Authority hereby declares, according to Article 17(2)(a), that **no international search report will be established** on the international application for the reasons indicated below.

1. ☐ The subject matter of the international application relates to:
 - a. ☐ scientific theories
 - b. ☐ mathematical theories
 - c. ☐ plant varieties
 - d. ☐ animal varieties
 - e. ☐ essentially biological processes for the production of plants and animals, other than microbiological processes and the products of such processes
 - f. ☐ schemes, rules or methods of doing business
 - g. ☐ schemes, rules or methods of performing purely mental acts
 - h. ☐ schemes, rules or methods of playing games
 - i. ☐ methods for treatment of the human body by surgery or therapy
 - j. ☐ methods for treatment of the animal body by surgery or therapy
 - k. ☐ diagnostic methods practised on the human or animal body
 - l. ☐ mere presentations of information
 - m. ☐ computer programs for which this International Searching Authority is not equipped to search prior art
2. ☐ The failure of the following parts of the international application to comply with prescribed requirements prevents a meaningful search from being carried out:

☐ the description
 ☐ the claims
 ☐ the drawings
3. ☒ A meaningful search could not be carried out without the sequence listing; the applicant did not, within the prescribed time limit:

☒ furnish a sequence listing in the form of an Annex C/ST.25 text file, and such listing was not available to the International Searching Authority in a form and manner acceptable to it; or the sequence listing furnished did not comply with the standard provided for in Annex C of the Administrative Instructions.

☐ furnish a sequence listing on paper or in the form of an image file complying with the standard provided for in Annex C of the Administrative Instructions, and such listing was not available to the International Searching Authority in a form and manner acceptable to it; or the sequence listing furnished did not comply with the standard provided for in Annex C of the Administrative Instructions.

☐ pay the required late furnishing fee for the furnishing of a sequence listing in response to an invitation under Rule 13^{ter}.1(a) or (b).
4. Further comments:

Applicant failed to submit a valid electronic seq. listing in response to the ISA/225.

| | |
|---|---|
| Name and mailing address of the ISA/ Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300 | Authorized officer <div style="text-align: right;">Blaine Copenheaver</div> <div style="font-size: small;"> PCT Helpdesk: 571-272-4300 PCT QSP: 571-272-7774 </div> |
|---|---|