



(51) International Patent Classification:

C12N 15/10 (2006.01) C07K 14/47 (2006.01)
A61K 48/00 (2006.01) C12N 15/113 (2010.01)

(21) International Application Number:

PCT/US2017/063468

(22) International Filing Date:

28 November 2017 (28.11.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/426,853 28 November 2016 (28.11.2016) US

(71) Applicant: THE BOARD OF REGENTS OF THE UNIVERSITY OF TEXAS SYSTEM [US/US]; 210 West 7th Street, Austin, TX 78701 (US).

(72) Inventors: ZHANG, Yu; c/o UT Southwestern Medical University, 5323 Harry Hines Blvd., Dallas, TX 75390 (US). LONG, Chengzu; c/o New York University School of Medicine, 522 First Avenue, Smilow 701C, New York, NY 10016 (US). BASSEL-DUBBY, Rhonda; 16106 Shadybank Drive, Dallas, TX 75248 (US). OLSON, Eric; 3219 Southwestern Blvd., University Park, TX 75225 (US).

(74) Agent: HIGHLANDER, Steven L.; Parker Highlander PLLC, 1120 S. Capital Of Texas Highway, Bldg. One, Suite 200, Austin, TX 78746 (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, JO, JP, KE, KG, KH, 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 international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))
- with sequence listing part of description (Rule 5.2(a))

(54) Title: PREVENTION OF MUSCULAR DYSTROPHY BY CRISPR/CPFI-MEDIATED GENE EDITING

(57) Abstract: Duchenne muscular dystrophy (DMD) is an inherited X-linked disease caused by mutations in the gene encoding dystrophin, a protein required for muscle fiber integrity. The disclosure reports CRISPR/Cpfi-mediated gene editing (Myo-editing) is effective at correcting the dystrophin gene mutation in the mdx mice, a model for DMD. Further, the disclosure reports optimization of germline editing of mdx mice by engineering the permanent skipping of mutant exon and extending exon skipping to also correct the disease by post-natal delivery of adeno-associated virus (AAV). AAV-mediated Myo-editing can efficiently rescue the reading frame of dystrophin in mdx mice in vivo. The disclosure reports means of Myo-editing-mediated exon skipping has been successfully advanced from somatic tissues in mice to human DMD patients-derived iPSCs (induced pluripotent stem cells). Custom Myo-editing was performed on iPSCs from patients with differing mutations and successfully restored dystrophin protein expression for all mutations in iPSCs-derived cardiomyocytes.



DESCRIPTION**PREVENTION OF MUSCULAR DYSTROPHY BY CRISPR/CPF1-MEDIATED
GENE EDITING****PRIORITY CLAIM**

This application claims priority to U.S. Provisional Patent Application Serial No. 62/426,853 which was filed on November 28, 2016, and entitled "Prevention of Muscular Dystrophy by CRISPR/Cpf1-Mediated Gene Editing," the disclosure of which is hereby incorporated by reference in its entirety.

FEDERAL FUNDING SUPPORT CLAUSE

This invention was made with government support under DK-099653 and U54-HD 087351 awarded by National Institutes of Health. The government has certain rights in the invention.

SEQUENCE LISTING

The instant application contains a Sequence Listing which has been submitted electronically in ASCII format and is hereby incorporated by reference in its entirety. Said ASCII copy, created on November 28, 2017, is named UTFD_3124WO.txt and is 189,059 bytes in size.

FIELD OF THE DISCLOSURE

The present disclosure relates to the fields of molecular biology, medicine and genetics. More particularly, the disclosure relates to the use of genome editing to treat Duchenne muscular dystrophy (DMD).

BACKGROUND

Duchenne muscular dystrophy (DMD) is an X-linked recessive disease caused by mutations in the gene coding for dystrophin, which is a large cytoskeletal protein essential for integrity of muscle cell membranes. DMD causes progressive muscle weakness, culminating

in premature death by the age of 30, generally from cardiomyopathy. There is no effective treatment for this disease. Numerous approaches to rescue dystrophin expression in DMD have been attempted, including delivery of truncated dystrophin or utrophin by recombinant adeno-associated virus (rAAV) and skipping of mutant exons with anti-sense
5 oligonucleotides and small molecules. However, these approaches cannot correct DMD mutations or permanently restore dystrophin expression. Accordingly, there is a need in the art for compositions and methods for treating DMD that correct DMD mutations to address the underlying cause of the disease, thereby permanently restore dystrophin expression.

SUMMARY

The disclosure provides a composition comprising a sequence encoding a Cpf1 polypeptide and a sequence encoding a DMD guide RNA (gRNA), wherein the DMD gRNA targets a dystrophin splice site, and wherein the DMD gRNA comprises any one of SEQ ID No. 448 to 770. In some embodiments, the sequence encoding the Cpf1 polypeptide is
15 isolated or derived from a sequence encoding a *Lachnospiraceae* Cpf1 polypeptide. In some embodiments, the sequence encoding the Cpf1 polypeptide is isolated or derived from a sequence encoding an *Acidaminococcus* Cpf1 polypeptide. In some embodiments, the sequence encoding the Cpf1 polypeptide or the sequence encoding the DMD gRNA comprises an RNA sequence. In some embodiments, the RNA sequence is an mRNA
20 sequence. In some embodiments, the RNA sequence comprises at least one chemically-modified nucleotide. In some embodiments, the sequence encoding the Cpf1 polypeptide comprises a DNA sequence.

In some embodiments, a first vector comprises the sequence encoding the Cpf1 polypeptide and a second vector comprises the sequence encoding the DMD gRNA. In some
25 embodiments, the first vector or the sequence encoding the Cpf1 polypeptide further comprises a first poly A sequence. In some embodiments, the second vector or the sequence encoding the DMD gRNA further comprises a second poly A sequence. In some embodiments, the first vector or the second vector further comprises a sequence encoding a detectable marker. In some embodiments, the detectable marker is a fluorescent maker.

In some embodiments, the first vector or the sequence encoding the Cpf1 polypeptide further comprises a first promoter sequence. In some embodiments, the second vector or the sequence encoding the DMD gRNA further comprises a second promoter sequence. In some
30 embodiments, the promoter first promoter sequence and the second promoter sequence are identical. In some embodiments, the first promoter sequence and the second promoter

sequence are not identical. In some embodiments, the first promoter sequence or the second promoter sequence comprises a constitutive promoter. In some embodiments, the first promoter sequence or the second promoter sequence comprises an inducible promoter. In some embodiments, the first promoter sequence or the second promoter sequence comprises a muscle-cell specific promoter. In some embodiments, the muscle-cell specific promoter is a myosin light chain-2 promoter, an α -actin promoter, a troponin 1 promoter, a $\text{Na}^+/\text{Ca}^{2+}$ exchanger promoter, a dystrophin promoter, an $\alpha 7$ integrin promoter, a brain natriuretic peptide promoter, an α B-crystallin/small heat shock protein promoter, an α -myosin heavy chain promoter, or an ANF promoter.

In some embodiments, the first vector or the second vector further comprises a sequence encoding 2A-like self-cleaving domain. In some embodiments, the sequence encoding 2A-like self-cleaving domain comprises a TaV-2A peptide.

In some embodiments, the vector comprises the sequence encoding the Cpf1 polypeptide and the sequence encoding the DMD gRNA. In embodiments, the vector further comprises a polyA sequence. In embodiments, the vector further comprises a promoter sequence. In embodiments, the promoter sequence comprises a constitutive promoter. In further embodiments, the promoter sequence comprises an inducible promoter. In embodiments, the promoter sequence comprises a muscle-cell specific promoter. In some embodiments, the muscle-cell specific promoter is a myosin light chain-2 promoter, an α -actin promoter, a troponin 1 promoter, a $\text{Na}^+/\text{Ca}^{2+}$ exchanger promoter, a dystrophin promoter, an $\alpha 7$ integrin promoter, a brain natriuretic peptide promoter, an α B-crystallin/small heat shock protein promoter, an α -myosin heavy chain promoter, or an ANF promoter.

In embodiments, the composition comprises a sequence codon optimized for expression in a mammalian cell. In further embodiments, the composition comprises a sequence codon optimized for expression in a human cell. In embodiments, the sequence encoding the Cpf1 polypeptide is codon optimized for expression in human cells.

In some embodiments, the splice site is a splice donor site. In some embodiments, the splice site is a splice acceptor site.

In further embodiments, the first vector or the second vector is a non-viral vector. In embodiments, the non-viral vector is a plasmid. In embodiments, a liposome or a nanoparticle comprises the first vector or the second vector.

In embodiments, the first vector or the second vector is a viral vector. In embodiments, the viral vector is an adeno-associated viral (AAV) vector. In embodiments,

the AAV vector is replication-defector or conditionally replication defective. In
embodiments, the AAV vector is a recombinant AAV vector. In some embodiments, the
AAV vector comprises a sequence isolated or derived from an AAV vector of serotype
AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 or
5 any combination thereof.

In some embodiments, the composition further comprises a single-stranded DMD
oligonucleotide. In some embodiments, the composition further comprises a
pharmaceutically acceptable carrier.

Also provided is a cell comprising a composition of the disclosure. In embodiments,
10 the cell is a muscle cell, a satellite cell or a precursor thereof. In some embodiments, the cell
is an iPSC or an iCM.

Also provided is a composition comprising a cell of the instant disclosure.

Also provided is a method of correcting a dystrophin gene defect comprising
contacting a cell and a composition of the disclosure under conditions suitable for expression
15 of the Cpf1 polypeptide and the gRNA, wherein the Cpf1 polypeptide disrupts the dystrophin
splice site; and wherein disruption of the splice site results in selective skipping of a mutant
DMD exon. In some embodiments, the mutant DMD exon is exon 23. In some
embodiments, the mutant DMD exon is exon 51. In embodiments, the cell is in vivo, ex vivo,
in vitro or in situ.

20 The disclosure also provides a of treating muscular dystrophy in a subject in need
thereof comprising administering to the subject a therapeutically effective amount of a
composition according to the instant disclosure. In embodiments, the composition is
administered locally. In embodiments, the composition is administered directly to a muscle
tissue. In embodiments, the composition is administered by intramuscular infusion or
25 injection. In embodiments, the muscle tissue comprises a tibialis anterior tissue, a quadricep
tissue, a soleus tissue, a diaphragm tissue or a heart tissue. In some embodiments, the
composition is administered systemically. In some embodiments, the composition is
administered by intravenous infusion or injection.

In embodiments, following administration of the composition, the subject exhibits
30 normal dystrophin-positive myofibers, and mosaic dystrophin-positive myofibers containing
centralized nuclei, or a combination thereof. In some embodiments, following administration
of the composition, the subject exhibits an emergence or an increase in a level of abundance
of normal dystrophin-positive myofibers when compared to an absence or a level of
abundance of normal dystrophin-positive myofibers prior to administration of the

composition. In some embodiments, following administration of the composition, the subject exhibits an emergence or an increase in a level of abundance of mosaic dystrophin-positive myofibers containing centralized nuclei when compared to an absence or an level of abundance of mosaic dystrophin-positive myofibers containing centralized nuclei prior to administration of the composition. In some embodiments, the subject exhibits a decreased serum CK level when compared to a serum CK level prior to administration of the composition. In embodiments, following administration of the composition, the subject exhibits improved grip strength when compared to a grip strength prior to administration of the composition.

In some embodiments, the method comprises administering a therapeutically effective amount of a composition disclosed herein, wherein the cell is autologous. In some embodiments, the method comprises administering a therapeutically effective amount of the composition, wherein the cell is allogeneic.

In embodiments, the subject is a neonate, an infant, a child, a young adult, or an adult.

In embodiments, the subject has muscular dystrophy. In embodiments, the subject is a genetic carrier for muscular dystrophy. In some embodiments, the subject is male. In some embodiments, the subject is female. In some embodiments, the subject appears to be asymptomatic and wherein a genetic diagnosis reveals a mutation in one or both copies of a *DMD* gene that impairs function of the *DMD* gene product. In some embodiments, the subject presents an early sign or symptom of muscular dystrophy. In some embodiments, the early sign or symptom of muscular dystrophy comprises loss of muscle mass or proximal muscle weakness. In embodiments, the loss of muscle mass or proximal muscle weakness occurs in one or both leg(s) and/or a pelvis, followed by one or more upper body muscle(s). In embodiments, the early sign or symptom of muscular dystrophy further comprises pseudohypertrophy, low endurance, difficulty standing, difficulty walking, and/or difficulty ascending a staircase or a combination thereof. In embodiments, the subject presents a progressive sign or symptom of muscular dystrophy. In embodiments, the progressive sign or symptom of muscular dystrophy comprises muscle tissue wasting, replacement of muscle tissue with fat, or replacement of muscle tissue with fibrotic tissue. In embodiments, the subject presents a later sign or symptom of muscular dystrophy. In some embodiments, the later sign or symptom of muscular dystrophy comprises abnormal bone development, curvature of the spine, loss of movement, and paralysis. In embodiments, the subject presents a neurological sign or symptom of muscular dystrophy. In embodiments, the neurological sign or symptom of muscular dystrophy comprises intellectual impairment and paralysis. In

embodiments, administration of the composition occurs prior to the subject presenting one or more progressive, later or neurological signs or symptoms of muscular dystrophy. In some embodiments, the subject is less than 10 years old. In some embodiments, the subject is less than 5 years old. In some embodiments, the subject is less than 2 years old.

5 The disclosure also provides a use of a therapeutically-effective amount of a composition for treating muscular dystrophy in a subject in need thereof.

It is contemplated that any method or composition described herein can be implemented with respect to any other method or composition described herein.

10 Other objects, features and advantages of the present disclosure will become apparent from the following detailed description. It should be understood, however, that the detailed description and the specific examples, while indicating specific embodiments of the disclosure, are given by way of illustration only, since various changes and modifications within the spirit and scope of the disclosure will become apparent to those skilled in the art from this detailed description.

15 **BRIEF DESCRIPTION OF THE DRAWINGS**

The following drawings form part of the present specification and are included to further demonstrate certain aspects of the present disclosure. The disclosure may be better understood by reference to one or more of these drawings in combination with the detailed description of specific embodiments presented herein.

20 **FIGS. 1A-E. Correction of DMD mutations by Cpf1-mediated genome editing.**
(FIG. 1A) A *DMD* deletion of exons 48-50 results in splicing of exon 47 to 51, generating an out-of-frame mutation of dystrophin. Two strategies were used for the restoration of dystrophin expression by Cpf1. In the “reframing” strategy, small INDELs in exon 51 restore the protein reading frame of dystrophin. The “exon skipping” strategy is achieved by
25 disruption of the splice acceptor of exon 51, which results in splicing of exon 47 to 52 and restoration of the protein reading frame. (FIG. 1B) The 3' end of an intron is T-rich, which generates Cpf1 PAM sequences enabling genome cleavage by Cpf1. (FIG. 1C) Illustration of Cpf1 gRNA targeting *DMD* exon 51. The T-rich PAM (red line) is located upstream of exon 51 near the splice acceptor site. The sequence of the Cpf1 g1 gRNA targeting exon 51 is
30 shown, highlighting the complementary nucleotides in blue. Cpf1 cleavage produces a staggered-end distal to the PAM site (demarcated by red arrowheads). The 5' region of exon 51 is shaded in light blue. Exon sequence is upper case. Intron sequence is lower case. (FIG. 1D) Illustration of a plasmid encoding human codon-optimized Cpf1 (hCpf1) with a nuclear

localization signal (NLS) and 2A-GFP. The plasmid also encodes a Cpf1 gRNA driven by the U6 promoter. Cells transfected with this plasmid express GFP, allowing for selection of Cpf1-expressing cells by FACS. (FIG. 1E) T7E1 assays using human 293T cells or DMD iPSCs (RIKEN51) transfected with plasmid expressing LbCpf1 or AsCpf1, gRNA and GFP show genome cleavage at *DMD* exon 51. Red arrowheads point to cleavage products. M, marker.

FIGS. 2A-I. DMD iPSC-derived cardiomyocytes express dystrophin after Cpf1-mediated genome editing by reframing. (FIG. 2A) DMD skin fibroblast-derived iPSCs were edited by Cpf1 using gRNA (corrected DMD-iPSCs) and then differentiated into cardiomyocytes (corrected cardiomyocytes) for analysis of genetic correction of the DMD mutation. (FIG. 2B) A *DMD* deletion of exons 48-50 results in splicing of exon 47 to 51, generating an out-of-frame mutation of dystrophin. Forward primer (F) targeting exon 47 and reverse primer (R) targeting exon 52 were used in RT-PCR to confirm the reframing strategy by Cpf1-mediated genome editing in cardiomyocytes. Uncorrected cardiomyocytes lack exons 48-50. In contrast, after reframing, exon 51 is placed back in-frame with exon 47. (FIG. 2C) Sequencing of representative RT-PCR products shows that uncorrected DMD iPSC-derived cardiomyocytes have a premature stop codon in exon 51, which creates a nonsense mutation. After Cpf1-mediated reframing, the ORF of dystrophin is restored. Dashed red line denotes exon boundary. (FIG. 2D) Western blot analysis shows dystrophin expression in a mixture of DMD iPSC-derived cardiomyocytes edited by reframing with LbCpf1 or AsLpf1 and g1 gRNA. Even without clonal selection, Cpf1-mediated reframing is efficient and sufficient to restore dystrophin expression in the cardiomyocyte mixture. α MHC is loading control. (FIG. 2E) Immunocytochemistry shows dystrophin expression in iPSC-derived cardiomyocyte (CM) mixtures following LbCpf1- or AsCpf1-mediated reframing. Dystrophin staining (red); Troponin I staining (green). Scale bar = 100 microns. (FIG. 2F) Western blot analysis shows dystrophin expression in single clones (#2 and #5) of iPSC-derived cardiomyocytes following clonal selection after LbCpf1-mediated reframing. α MHC is loading control. (FIG. 2G) Immunocytochemistry showing dystrophin expression in clone #2 LbCpf1-edited iPSC-derived cardiomyocytes. Scale bar = 100 microns. (FIG. 2H) Quantification of mitochondrial DNA copy number in single clones (#2 and #5) of LbCpf1-edited iPSC-derived cardiomyocytes. Data are represented as mean $\pm$ SEM (n = 3). (&) P < 0.01; (#) P < 0.005; (ns) not significant. (FIG. 2I) Basal oxygen consumption rate (OCR) of single clones (#2 and #5) of LbCpf1-edited iPSC-derived cardiomyocytes, and OCR in response to oligomycin, FCCP, and Rotenone and Antimycin A, normalized to cell number

(order left to right for each test is the same as FIG. 2H). Data are represented as mean $\pm$ SEM (n = 5). (*) P < 0.05; (&) P < 0.01; (#) P < 0.005; (ns) not significant.

FIGS. 3A-H. DMD iPSC-derived cardiomyocytes express dystrophin after Cpf1-mediated exon skipping. (FIG. 3A) Two gRNAs, either gRNA (g2 or g3), which target

5 intron 50, and the other (g1), which targets exon 51, were used to direct Cpf1-mediated removal of the exon 51 splice acceptor site. (FIG. 3B) T7E1 assay using 293T cells transfected with LbCpf1 and gRNA2 (g2) or gRNA3 (g3) shows cleavage of the DMD locus at intron 50. Red arrowheads denote cleavage products. M, marker. (FIG. 3C) PCR products of genomic DNA isolated from DMD-iPSCs transfected with a plasmid expressing
10 LbCpf1, g1 + g2 and GFP. The lower band (red arrowhead) indicates removal of the exon 51 splice acceptor site. (FIG. 3D) Sequence of the lower PCR band from panel c shows a 200-bp deletion, spanning from the 3'-end of intron 50 to the 5'-end of exon 51. This confirms removal of the "ag" splice acceptor of exon 51. The sequence of the uncorrected allele is shown above that of the LbCpf1-edited allele. (FIG. 3E) RT-PCR of iPSC-derived
15 cardiomyocytes using primer sets described in Fig. 2B. The 700-bp band in the WT lane is the dystrophin transcript from exon 47-52; the 300-bp band in the uncorrected lane is the dystrophin transcript from exon 47-52 with exon 48-50 deletion; and the lower band in the g1+g2 mixture lane (edited by LbCpf1) shows exon 51 skipping. (FIG. 3F) Sequence of the lower band from panel e (g1+g2 mixture lane) confirms skipping of exon 51, which reframed
20 the DMD ORF. (FIG. 3G) Western blot analysis shows dystrophin protein expression in iPSC-derived cardiomyocyte mixtures after exon 51 skipping by LbCpf1 with g1 + g2. α MHC is loading control. (FIG. 3H) Immunocytochemistry shows dystrophin expression in iPSC-derived cardiomyocyte mixtures (CMs) following Cpf1-mediated exon skipping with g1 + g2 gRNA compared to WT and uncorrected CMs. Dystrophin staining (red). Troponin I
25 staining (green). Scale bar = 100 microns.

FIGS. 4A-D. CRISPR-Cpf1-mediated editing of exon 23 of the mouse DMD gene.

(FIG. 4A) Illustration of mouse *Dmd* locus highlighting the mutation at exon 23. Sequence shows the nonsense mutation caused by C to T transition, which creates a premature stop codon. (FIG. 4B) Illustration showing the targeting location of gRNAs (g1, g2 and g3)
30 (shown in light blue) on exon 23 of the *Dmd* gene. Red line represents LbCpf1 PAM. (FIG. 4C) T7E1 assay using mouse 10T1/2 cells transfected with LbCpf1 or AsCpf1 with different gRNAs (g1, g2 or g3) targeting exon 23 shows that LbCpf1 and AsCpf1 have different cleavage efficiency at the *Dmd* exon 23 locus. Red arrowheads show cleavage products of genome editing. M, marker. (FIG. 4D) Illustration of LbCpf1-mediated gRNA (g2) targeting

of *Dmd* exon 23. Red arrowheads indicate the cleavage site. The ssODN HDR template contains the *mdx* correction, four silent mutations (green) and a *TseI* restriction site (underlined).

FIGS. 5A-F. CRISPR-LbCpf1-mediated *Dmd* correction in *mdx* mice. (FIG. 5A)

5 Strategy of gene correction in *mdx* mice by LbCpf1-mediated germline editing. Zygotes from intercrosses of *mdx* parents were injected with gene editing components (LbCpf1 mRNA, g2 gRNA and ssODN) and reimplanted into pseudo-pregnant mothers, which gave rise to pups with gene correction (*mdx-C*). (FIG. 5B) Illustration showing LbCpf1 correction of *mdx* allele by HDR or NHEJ. (FIG. 5C) Genotyping results of LbCpf1-edited *mdx* mice. Top
10 panel shows T7E1 assay. Blue arrowhead denotes uncleaved DNA and red arrowhead shows T7E1 cleaved DNA. Bottom panel shows *TseI* RFLP assay. Blue arrowhead denotes uncorrected DNA. Red arrowhead points to *TseI* cleavage indicating HDR correction. *mdx-C*1-C5 denotes LbCpf1-edited *mdx* mice. (FIG. 5D) Top panel shows sequence of WT *Dmd* exon 23. Middle panel shows sequence of *mdx* *Dmd* exon 23 with C to T mutation, which
15 generates a STOP codon. Bottom panel shows sequence of *Dmd* exon 23 with HDR correction by LbCpf1-mediated editing. Black arrow points to silent mutations introduced by the ssODN HDR template. (FIG. 5E) H&E of tibialis anterior (TA) and gastrocnemius/plantaris (G/P) muscles from WT, *mdx* and LbCpf1-edited mice (*mdx-C*). (FIG. 5F) Immunohistochemistry of TA and G/P muscles from WT, *mdx* and LbCpf1-edited
20 mice (*mdx-C*) using antibody to dystrophin (red). *mdx* muscle showed fibrosis and inflammatory infiltration, whereas *mdx-C* muscle showed normal muscle structure.

FIGS. 6A-C. Genome editing at DMD exon 51 by LbCpf1 or AsCpf1. (FIG. 6A)

DNA sequencing of *DMD* exon 51 from a mixture of DMD patient (RIKEN 51) skin fibroblast-derived iPSCs edited by LbCpf1 or AsCpf1 using g1. Sequences of individual
25 edited DMD allele are shown beneath the uncorrected DMD allele. Δ denotes nucleotide deletion. (FIG. 6B) DNA sequencing of *DMD* exon 51 from a single clone of DMD patient skin fibroblast-derived iPSCs edited by LbCpf1 or AsCpf1 using g1. (FIG. 6C) DNA Sequencing of PCR products of 10T1/2 cells following LbCpf1-editing with g2 or g3. WT sequence is on top and INDEL sequences are on the bottom.

30 **FIGS. 7A-B. Histological analysis of muscles from WT, *mdx* and LbCpf1-edited mice (*mdx-C*).** (FIG. 7A) Immunohistochemistry and H&E staining of whole tibialis anterior (TA) muscle. Dystrophin staining is red. (FIG. 7B) Immunohistochemistry and H&E staining of whole gastrocnemius/plantaris (G/P) muscles. Dystrophin staining is red.

DETAILED DESCRIPTION

Duchenne muscular dystrophy, like many other diseases of genetic origin, present challenging therapeutic scenarios. The CRISPR-Cas system represents an approach for correction of diverse genetic defects. The CRISPR (clustered regularly interspaced short palindromic repeats) system functions as an adaptive immune system in bacteria and archaea that defends against phage infection. In this system, an endonuclease is guided to specific genomic sequences by a single guide RNA (sgRNA), resulting in DNA cutting near a protospacer adjacent motif (PAM) sequence. Previously, CRISPR-Cas9 was used to correct the DMD mutation in mice and human cells. However, many challenges remain to be addressed. For example, *Streptococcus pyogenes* Cas9 (SpCas9), currently the most widely used Cas9 endonuclease, has a G-rich PAM requirement (NGG) that excludes genome editing of AT-rich regions. Additionally, the large size of SpCas9 reduces the efficiency of packaging and delivery in low-capacity viral vectors, such as Adeno-associated virus (AAV) vectors. The Cas9 endonuclease from *Staphylococcus aureus* (SaCas9), although smaller in size than SpCas9, has a PAM sequence (NNGRRT) that is longer and more complex, thus limiting the range of its genomic targets (Ran *et al.*, 2015). Smaller CRISPR enzymes with greater flexibility in recognition sequence and comparable cutting efficiency would facilitate precision gene editing, especially for translational applications.

As demonstrated by the disclosure, an RNA-guided endonuclease, named Cpf1 (CRISPR from *Prevotella* and *Francisella* 1), is effective for mammalian genome cleavage. Cpf1 has several unique features that expand its genome editing potential when compared to Cas9: Cpf1-mediated cleavage is guided by a single and short crRNA (abbreviated as gRNA), whereas Cas9-mediated cleavage is guided by a hybrid of CRISPR RNA (crRNA) and a long trans-activating crRNA (tracrRNA). Cpf1 prefers a T-rich PAM at the 5'-end of a protospacer, while Cas9 requires a G-rich PAM at the 3' end of the target sequence. Cpf1-mediated cleavage produces a sticky end distal to the PAM site, which activates DNA repair machinery, while Cas9 cutting generates a blunt end. Cpf1 also has RNase activity, which can process precursor crRNAs to mature crRNAs. Like Cas9, Cpf1 binds to a targeted genomic site and generates a double-stranded break (DSB), which is then repaired either by non-homologous end-joining (NHEJ) or by homology-directed repair (HDR) if an exogenous template is provided. Prior to the instant disclosure, neither had the advantages of Cpf1 over Cas9 been appreciated nor had the use of Cpf1 for correction of genetic mutations in mammalian cells and animal models of disease been demonstrated. Here, the inventors show

that Cpf1 provides a robust and efficient RNA-guided genome editing system that permanently corrects DMD mutations by different strategies, thereby restoring dystrophin expression and preventing progression of the disease. These findings provide a new approach for the permanent correction of human genetic mutations. These and other aspects of the disclosure are reproduced below.

I. Duchenne Muscular Dystrophy

A. Background

Duchenne muscular dystrophy (DMD) is a recessive X-linked form of muscular dystrophy, affecting around 1 in 3,500 boys, which results in muscle degeneration and premature death. The disorder is caused by a mutation in the gene dystrophin (See GenBank Accession No. NC_000023.11), located on the human X chromosome, which codes for the protein dystrophin (GenBank Accession No. AAA53189; SEQ ID NO. 383), the sequence of which is reproduced below:

```

15      1 mlwweevedc yeredvqkkt ftkwvnaqfs kfgkqhienl fsdlqdgrrl ldlllegltgq
      61 klpkckgstr vhalnnvnka lrvlqnnnvd lvnigstdiv dgnhkltlgl iwniilhwqv
      121 knvmknimag lqqttnsekil lswvrqstrn ypqvnvinft tswsdglaln alihshrpdl
      181 fdwnsvvcqg satqrlehaf niaryqlgie klldpedvdt typdkksilm yitslfqvlp
      241 qqvsieaiqe vemlprppkv tkeehfqlhh qmhysqqitv slaqgyerts spkprfksya
20      301 ytqaayvtts dptrspfpsq hleapedksf gsslmesevn ldryqtalee vlswlisaed
      361 tlqaqgeisn dvevvkdqfh thegymmdlt ahqgrvgnil qlgskligtg klseedeetev
      421 qeqmnlinsr weclrvasme kqsnlhrvlm dlqnqklkel ndwltkter trkmeeeplg
      481 pdledlkrqv qqhkvqlqedl egeqvrnsi thmvvvvdes sgdhataale eqklvlgdrw
      541 anicrwtedr wvllqdillk wqrlteeqcl fsawlseked avnkihttgf kdqnemlssl
      601 qklavlkadl ekkkqsmgkl yslkqdlst lknksvtqkt eawldnfarc wdnlvqklek
25      661 staqisqavt ttqpsltqtt vmetvttvtt reqilvkhaq eelpppppqk krqitvdsei
      721 rkrlldvdite lhwitrsea vlqspafaif rkegnfsdlk ekvnaierk aekfrklqda
      781 srsaqalveg mnvegvnads ikqaseqlns rwieefcqls erlnwleyqn niiafynqlq
      841 qlegmtttae nwlkiqpttp septaiksqr kickdevnrl sglqpqierl kqisialkek
      901 gggpmfldad fvaftnhfkq vdsdvqarek elqtifdtlp pmryqetmsa irtwvqqset
30      961 klsipqlsvt dyeimeqrlg elqalqsslg eqqsglyyls ttvkemskka pseisrkyqs
      1021 efeeiegrwk klssqlvehc qkleeqmnkl rkighniqtl kkwmaevdvf lkeewpalgd
      1081 seilkklqkq crllvsdiqt iqpslnsvne ggqkikneae pefasrlete lkelntqwdh
      1141 mcqqvyarke alkGGLEktv slqkdlsesm ewmtqaeey lerdfeyktp delqkaveem
      1201 krakeeaaqk eakvklites vnsviaqapp vagealkkel etltnnyqwl ctrlngkckt
35      1261 leevwacwhe llsylekank wlnevefklk ttenipggae eisevldsle nlmrhsednp
      1321 nqirilaqtl tdggvmdeli neeletfnr wrclheevr rqlleqsiq saqetekslh
      1381 lqesltfid kqlaayiadk vdaaqmpqea qkiqsdlts eisleemkqh nqgkeaaqrv
      1441 lsqidvaqkk lqdvsmkfrl fqkpanfelr lqeskmilde vkmhlpalet ksveqevvqs
      1501 qlnhcvnlyk slsevkseve mviktgrqiv qkkqtenpke ldervtalkl hynelgakvt
40      1561 erkqglekcl klsrkmrkem nvltewlaat dmeltkrsav egmpsnldse vawgkatqke
      1621 iekqkvhlks itevgealkt vlqketlve dklslslsnw iavtsraeew lnlleyqkh
      1681 metfdqnvdh itkwiiqadt lldesekkkp qqkedvlkrl kaelndirpk vdstrdqaan
      1741 lmanrgdhcr klvepqisel nhrfaaishr igtgkasipl keleqfnssi qkllepleae
      1801 iqggvnklee dfnkdmnedn egtvkellqr gdnllqgritd erkreeikik qqlqltkhna
45      1861 lkdlrsqrrk kaleishqwy qykrqaddll kclddiekkl aslpeprder kikeidrelq
      1921 kkeeklnavr rqaeglsedg aamaveptqi qlskrwreie skfaqfrrln faqihvree
      1981 tmmvmtedmp leisyvpsty lteithvsqa lleveqllna pdlcakdfed lfkqeeslkn
      2041 ikdslqqssg ridihskkk aalqsatpve rvklqealsq ldfqwekvnk mykdrqgrfd
      2101 rsvekwrfrh ydikifnqwl teaeqflrkt qipenwehak ykwylkelqd gigqrqtvvr

```

2161 tlnatgeeii qqssktdasi lqeklgslnl rwqevckqls drkkrleeqk nilsefqrld
 2221 nefvlwleea dniasiplep gkeqqlkekl eqvklleveel plrqgilkql netggpvlvs
 2281 apispeeek lenklkqtnl qwikvsralp ekqgeieaqi kdlgglekkl edleeqlnhl
 2341 llwlsplrnl leiynqpnqe gpfdvqetei avqakqpdve eilskgqhly kekpatqpvk
 5 2401 rkledlssew kavnrllqel rakqpdlapg lttigasptq tvtlvtqpvv tketaiskle
 2461 mpsslmlevp aladfnrawt eltdwlsld qviksqrvmv gdledinemi ikqkatmqdl
 2521 eqrrpqleel itaaqnlknk tsnqeartii tdrieriqnq wdevqehlqn rrqqlnemlk
 2581 dstqwleake eaeqvlqgar akleswkegp ytvdaikqi tetkqlakdl rqwqtnvdva
 10 2641 ndlakllrd ysaddtrkvh miteninasw rsihkrvser eaaleethrl lqqfpldle
 2701 flawlteaet tanvlqdatr kerlledskg vkelmkqwqd lggeieahtd vyhnl densq
 2761 kilrslegsd davllqrrld nmnfkwseir kkslnirshl eassdqwkrl hlsllqellvw
 2821 lqlkddelsr qapiggdfpa vqkqndvhra fkrelktkep vimstletvr iflteqpleg
 2881 leklyqepre lppeeraqnv trllrkqae vnteweklnl hsadwqrkid etlerlqelq
 15 2941 eatdeldlkl rqaevikgsw qpvgdlld lqdhlekvka lrgeiaplke nvshvndlar
 3001 qlttlgiqls pynlstledl ntrwkllqva vedrvrqlhe ahrdfgpasq hflstsvqgp
 3061 weraispnkvy ppyinhettc tcwdhpkmtc lyqsladlnn vrfsayrtam klrrllqkalc
 3121 ldllslsaac daldqhnkq ndqpm dilqi inclt tiydr leqehnnlvn vplcvdmcln
 3181 wllnvdytgr tgrirvlsfk tgiislckah ledkyrylfk qvasstgfc d qrrlgllldh
 20 3241 siqiprqlge vasfggsnie psvrscfqa nnkpeieaal fldwmrlepq smvwlpvlhr
 3301 vaaetaakhq akcnickecp iigfryrslk hfnydicqsc ffsgrvakgh kmhymveyc
 3361 tpttsgeivr dfakvlknkf rtkryfakhp rmgyllpvqtv legdnmetpv tlinfwpvds
 3421 apasspqlsh ddthsrlehy asrlaemens ngsylndsis pnesiddehl liqhycqsln
 3481 qdsplsqprs paqilisles eergeleril adleeeenrl qaeydrllkq hehkglsppl
 3541 sppemmtsp qsprdaelia eakllrqhkg rlearnqile dhnkqlsqsl hrlrlleqp
 25 3601 qaeakvngtt vsspstslqr sdssqpmllr vvgstsdsm geedllsppq dtstgleevm
 3661 eqlnnsfpss rgrntpgkpm redtm

In humans, dystrophin mRNA contains 79 exons. Dystrophin mRNA is known to be
 alternatively spliced, resulting in various isoforms. Exemplary dystrophin isoforms are listed
 30 in Table 1.

In humans, dystrophin mRNA contains 79 exons. Dystrophin mRNA is known to be
 alternatively spliced, resulting in various isoforms of the protein. Exemplary dystrophin
 isoforms are listed in Table 1.

35 Table 1: Dystrophin isoforms

Sequence Name*	Nucleic Acid Accession No.*	Nucleic Acid SEQ ID NO:	Protein Accession No.*	Protein SEQ ID NO:	Description
DMD Genomic Sequence	NC_000023.11 (positions 31119219 to 33339609)	None	None	None	Sequence from Human X Chromosome (at positions Xp21.2 to p21.1) from Assembly GRCh38.p7 (GCF_000001405.33)
Dystrophin Dp427c isoform	NM_000109.3	384	NP_000100.2	385	Transcript Variant: transcript Dp427c is expressed predominantly in neurons of the cortex and the CA regions of the hippocampus. It uses a unique

Sequence Name*	Nucleic Acid Accession No.*	Nucleic Acid SEQ ID NO:	Protein Accession No.*	Protein SEQ ID NO:	Description
					promoter/exon 1 located about 130 kb upstream of the Dp427m transcript promoter. The transcript includes the common exon 2 of transcript Dp427m and has a similar length of 14 kb. The Dp427c isoform contains a unique N-terminal MED sequence, instead of the MLWWEVEDCY (SEQ ID NO: 3) sequence of isoform Dp427m. The remainder of isoform Dp427c is identical to isoform Dp427m.
Dystrophin Dp427m isoform	NM_004006.2	386	NP_003997.1	387	Transcript Variant: transcript Dp427m encodes the main dystrophin protein found in muscle. As a result of alternative promoter use, exon 1 encodes a unique N-terminal MLWWEVEDCY (SEQ ID NO: 3) aa sequence.
Dystrophin Dp427p1 isoform	NM_004009.3	388	NP_004000.1	389	Transcript Variant: transcript Dp427p1 initiates from a unique promoter/exon 1 located in what corresponds to the first intron of transcript Dp427m. The transcript adds the common exon 2 of Dp427m and has a similar length (14 kb). The Dp427p1 isoform replaces the MLWWEVEDCY (SEQ ID NO: 3) -start of Dp427m with a unique N-terminal MSEVSSD (SEQ ID NO: 8) aa sequence.
Dystrophin Dp260-1 isoform	NM_004011.3	390	NP_004002.2	391	Transcript Variant: transcript Dp260-1 uses exons 30-79, and

Sequence Name*	Nucleic Acid Accession No.*	Nucleic Acid SEQ ID NO:	Protein Accession No.*	Protein SEQ ID NO:	Description
					originates from a promoter/exon 1 sequence located in intron 29 of the dystrophin gene. As a result, Dp260-1 contains a 95 bp exon 1 encoding a unique N-terminal 16 aa MTEIILLIFFPAYFLN-sequence that replaces amino acids 1-1357 of the full-length dystrophin product (Dp427m isoform).
Dystrophin Dp260-2 isoform	NM_004012.3	392	NP_004003.1	393	Transcript Variant: transcript Dp260-2 uses exons 30-79, starting from a promoter/exon 1 sequence located in intron 29 of the dystrophin gene that is alternatively spliced and lacks N-terminal amino acids 1-1357 of the full length dystrophin (Dp427m isoform). The Dp260-2 transcript encodes a unique N-terminal MSARKLRNLSYKK sequence.
Dystrophin Dp140 isoform	NM_004013.2	394	NP_004004.1	395	Transcript Variant: Dp140 transcripts use exons 45-79, starting at a promoter/exon 1 located in intron 44. Dp140 transcripts have a long (1 kb) 5' UTR since translation is initiated in exon 51 (corresponding to aa 2461 of dystrophin). In addition to the alternative promoter and exon 1, differential splicing of exons 71-74 and 78 produces at least five Dp140 isoforms. Of these, this transcript (Dp140) contains all of the exons.
Dystrophin Dp116 isoform	NM_004014.2	396	NP_004005.1	397	Transcript Variant: transcript Dp116 uses exons 56-79, starting from a promoter/exon 1

Sequence Name*	Nucleic Acid Accession No.*	Nucleic Acid SEQ ID NO:	Protein Accession No.*	Protein SEQ ID NO:	Description
					within intron 55. As a result, the Dp116 isoform contains a unique N-terminal MLHRKTYHVK aa sequence, instead of aa 1-2739 of dystrophin. Differential splicing produces several Dp116-subtypes. The Dp116 isoform is also known as S-dystrophin or apo-dystrophin-2.
Dystrophin Dp71 isoform	NM_004015.2	398	NP_004006.1	399	Transcript Variant: Dp71 transcripts use exons 63-79 with a novel 80- to 100-nt exon containing an ATG start site for a new coding sequence of 17 nt. The short coding sequence is in-frame with the consecutive dystrophin sequence from exon 63. Differential splicing of exons 71 and 78 produces at least four Dp71 isoforms. Of these, this transcript (Dp71) includes both exons 71 and 78.
Dystrophin Dp71b isoform	NM_004016.2	400	NP_004007.1	401	Transcript Variant: Dp71 transcripts use exons 63-79 with a novel 80- to 100-nt exon containing an ATG start site for a new coding sequence of 17 nt. The short coding sequence is in-frame with the consecutive dystrophin sequence from exon 63. Differential splicing of exons 71 and 78 produces at least four Dp71 isoforms. Of these, this transcript (Dp71b) lacks exon 78 and encodes a protein with a different C-terminus than Dp71 and Dp71a isoforms.
Dystrophin Dp71a isoform	NM_004017.2	402	NP_004008.1	403	Transcript Variant: Dp71 transcripts use

Sequence Name*	Nucleic Acid Accession No.*	Nucleic Acid SEQ ID NO:	Protein Accession No.*	Protein SEQ ID NO:	Description
					exons 63-79 with a novel 80- to 100-nt exon containing an ATG start site for a new coding sequence of 17 nt. The short coding sequence is in-frame with the consecutive dystrophin sequence from exon 63. Differential splicing of exons 71 and 78 produces at least four Dp71 isoforms. Of these, this transcript (Dp71a) lacks exon 71.
Dystrophin Dp71ab isoform	NM_004018.2	404	NP_004009.1	405	Transcript Variant: Dp71 transcripts use exons 63-79 with a novel 80- to 100-nt exon containing an ATG start site for a new coding sequence of 17 nt. The short coding sequence is in-frame with the consecutive dystrophin sequence from exon 63. Differential splicing of exons 71 and 78 produces at least four Dp71 isoforms. Of these, this transcript (Dp71ab) lacks both exons 71 and 78 and encodes a protein with a C-terminus like isoform Dp71b.
Dystrophin Dp40 isoform	NM_004019.2	406	NP_004010.1	407	Transcript Variant: transcript Dp40 uses exons 63-70. The 5' UTR and encoded first 7 aa are identical to that in transcript Dp71, but the stop codon lies at the splice junction of the exon/intron 70. The 3' UTR includes nt from intron 70 which includes an alternative polyadenylation site. The Dp40 isoform lacks the normal C-terminal end of full-length dystrophin (aa 3409-3685).

Sequence Name*	Nucleic Acid Accession No.*	Nucleic Acid SEQ ID NO:	Protein Accession No.*	Protein SEQ ID NO:	Description
Dystrophin Dp140c isoform	NM_004020.3	408	NP_004011.2	409	Transcript Variant: Dp140 transcripts use exons 45-79, starting at a promoter/exon 1 located in intron 44. Dp140 transcripts have a long (1 kb) 5' UTR since translation is initiated in exon 51 (corresponding to aa 2461 of dystrophin). In addition to the alternative promoter and exon 1, differential splicing of exons 71-74 and 78 produces at least five Dp140 isoforms. Of these, this transcript (Dp140c) lacks exons 71-74.
Dystrophin Dp140b isoform	NM_004021.2	410	NP_004012.1	411	Transcript Variant: Dp140 transcripts use exons 45-79, starting at a promoter/exon 1 located in intron 44. Dp140 transcripts have a long (1 kb) 5' UTR since translation is initiated in exon 51 (corresponding to aa 2461 of dystrophin). In addition to the alternative promoter and exon 1, differential splicing of exons 71-74 and 78 produces at least five Dp140 isoforms. Of these, this transcript (Dp140b) lacks exon 78 and encodes a protein with a unique C-terminus.
Dystrophin Dp140ab isoform	NM_004022.2	412	NP_004013.1	413	Transcript Variant: Dp140 transcripts use exons 45-79, starting at a promoter/exon 1 located in intron 44. Dp140 transcripts have a long (1 kb) 5' UTR since translation is initiated in exon 51 (corresponding to aa 2461 of dystrophin). In addition to the alternative promoter and exon 1, differential

Sequence Name*	Nucleic Acid Accession No.*	Nucleic Acid SEQ ID NO:	Protein Accession No.*	Protein SEQ ID NO:	Description
					splicing of exons 71-74 and 78 produces at least five Dp140 isoforms. Of these, this transcript (Dp140ab) lacks exons 71 and 78 and encodes a protein with a unique C-terminus.
Dystrophin Dp140bc isoform	NM_004023.2	414	NP_004014.1	415	Transcript Variant: Dp140 transcripts use exons 45-79, starting at a promoter/exon 1 located in intron 44. Dp140 transcripts have a long (1 kb) 5' UTR since translation is initiated in exon 51 (corresponding to aa 2461 of dystrophin). In addition to the alternative promoter and exon 1, differential splicing of exons 71-74 and 78 produces at least five Dp140 isoforms. Of these, this transcript (Dp140bc) lacks exons 71-74 and 78 and encodes a protein with a unique C-terminus.
Dystrophin isoform X2	XM_006724469.3	416	XP_006724532.1	417	
Dystrophin isoform X5	XM_011545467.1	418	XP_011543769.1	419	
Dystrophin isoform X6	XM_006724473.2	420	XP_006724536.1	421	
Dystrophin isoform X8	XM_006724475.2	422	XP_006724538.1	423	
Dystrophin isoform X4	XM_017029328.1	424	XP_016884817.1	425	
Dystrophin isoform X1	XM_006724468.2	426	XP_006724531.1	427	
Dystrophin isoform X13	XM_017029331.1	428	XP_016884820.1	429	
Dystrophin isoform X3	XM_006724470.3	430	XP_006724533.1	431	
Dystrophin isoform X7	XM_006724474.3	432	XP_006724537.1	433	
Dystrophin isoform X9	XM_011545468.2	434	XP_011543770.1	435	
Dystrophin isoform X11	XM_017029330.1	436	XP_016884819.1	437	
Dystrophin isoform X10	XM_017029329.1	438	XP_016884818.1	439	

Sequence Name*	Nucleic Acid Accession No.*	Nucleic Acid SEQ ID NO:	Protein Accession No.*	Protein SEQ ID NO:	Description
Dystrophin isoform X12	XM_011545469.1	440	XP_011543771.1	441	

Dystrophin is an important component within muscle tissue that provides structural stability to the dystroglycan complex (DGC) of the cell membrane. While both sexes can carry the mutation, females are rarely affected with the skeletal muscle form of the disease.

5

B. Symptoms

Symptoms usually appear in boys between the ages of 2 and 3 and may be visible in early infancy. Even though symptoms do not appear until early infancy, laboratory testing can identify children who carry the active mutation at birth. Progressive proximal muscle weakness of the legs and pelvis associated with loss of muscle mass is observed first. Eventually this weakness spreads to the arms, neck, and other areas. Early signs may include pseudohypertrophy (enlargement of calf and deltoid muscles), low endurance, and difficulties in standing unaided or inability to ascend staircases. As the condition progresses, muscle tissue experiences wasting and is eventually replaced by fat and fibrotic tissue (fibrosis). By age 10, braces may be required to aid in walking but most patients are wheelchair dependent by age 12. Later symptoms may include abnormal bone development that lead to skeletal deformities, including curvature of the spine. Due to progressive deterioration of muscle, loss of movement occurs, eventually leading to paralysis. Intellectual impairment may or may not be present but if present, does not progressively worsen as the child ages. The average life expectancy for males afflicted with DMD is around 25.

20

The main symptom of Duchenne muscular dystrophy, a progressive neuromuscular disorder, is muscle weakness associated with muscle wasting with the voluntary muscles being first affected, especially those of the hips, pelvic area, thighs, shoulders, and calves. Muscle weakness also occurs later, in the arms, neck, and other areas. Calves are often enlarged. Symptoms usually appear before age 6 and may appear in early infancy. Other physical symptoms are:

25

- Awkward manner of walking, stepping, or running – (patients tend to walk on their forefeet, because of an increased calf muscle tone. Also, toe walking is a compensatory adaptation to knee extensor weakness.)
- Frequent falls

30

- Fatigue
- Difficulty with motor skills (running, hopping, jumping)
- Lumbar hyperlordosis, possibly leading to shortening of the hip-flexor muscles. This has an effect on overall posture and a manner of walking, stepping, or running.
- 5 • Muscle contractures of Achilles tendon and hamstrings impair functionality because the muscle fibers shorten and fibrose in connective tissue
- Progressive difficulty walking
- Muscle fiber deformities
- Pseudohypertrophy (enlarging) of tongue and calf muscles. The muscle tissue is eventually replaced by fat and connective tissue, hence the term pseudohypertrophy.
- 10 • Higher risk of neurobehavioral disorders (*e.g.*, ADHD), learning disorders (dyslexia), and non-progressive weaknesses in specific cognitive skills (in particular short-term verbal memory), which are believed to be the result of absent or dysfunctional dystrophin in the brain
- 15 • Eventual loss of ability to walk (usually by the age of 12)
- Skeletal deformities (including scoliosis in some cases)
- Trouble getting up from lying or sitting position

The condition can often be observed clinically from the moment the patient takes his first steps, and the ability to walk usually completely disintegrates between the time the boy is 9 to 12 years of age. Most men affected with DMD become essentially “paralyzed from the neck down” by the age of 21. Muscle wasting begins in the legs and pelvis, then progresses to the muscles of the shoulders and neck, followed by loss of arm muscles and respiratory muscles. Calf muscle enlargement (pseudohypertrophy) is quite obvious. Cardiomyopathy particularly (dilated cardiomyopathy) is common, but the development of congestive heart failure or arrhythmia (irregular heartbeat) is only occasional.

A positive Gowers' sign reflects the more severe impairment of the lower extremities muscles. The child helps himself to get up with upper extremities: first by rising to stand on his arms and knees, and then “walking” his hands up his legs to stand upright. Affected children usually tire more easily and have less overall strength than their peers. Creatine kinase (CPK-MM) levels in the bloodstream are extremely high. An electromyography (EMG) shows that weakness is caused by destruction of muscle tissue rather than by damage to nerves. Genetic testing can reveal genetic errors in the Xp21 gene. A muscle biopsy

(immunohistochemistry or immunoblotting) or genetic test (blood test) confirms the absence of dystrophin, although improvements in genetic testing often make this unnecessary.

- Abnormal heart muscle (cardiomyopathy)
- Congestive heart failure or irregular heart rhythm (arrhythmia)
- 5 • Deformities of the chest and back (scoliosis)
- Enlarged muscles of the calves, buttocks, and shoulders (around age 4 or 5). These muscles are eventually replaced by fat and connective tissue (pseudohypertrophy).
- Loss of muscle mass (atrophy)
- Muscle contractures in the heels, legs
- 10 • Muscle deformities
- Respiratory disorders, including pneumonia and swallowing with food or fluid passing into the lungs (in late stages of the disease)

C. Causes

Duchenne muscular dystrophy (DMD) is caused by a mutation of the dystrophin gene at locus Xp21, located on the short arm of the X chromosome. Dystrophin is responsible for connecting the cytoskeleton of each muscle fiber to the underlying basal lamina (extracellular matrix), through a protein complex containing many subunits. The absence of dystrophin permits excess calcium to penetrate the sarcolemma (the cell membrane). Alterations in calcium and signaling pathways cause water to enter into the mitochondria, which then burst.

20 In skeletal muscle dystrophy, mitochondrial dysfunction gives rise to an amplification of stress-induced cytosolic calcium signals and an amplification of stress-induced reactive-oxygen species (ROS) production. In a complex cascading process that involves several pathways and is not clearly understood, increased oxidative stress within the cell damages the sarcolemma and eventually results in the death of the cell. Muscle fibers undergo necrosis and are ultimately replaced with adipose and connective tissue.

25 DMD is inherited in an X-linked recessive pattern. Females will typically be carriers for the disease while males will be affected. Typically, a female carrier will be unaware they carry a mutation until they have an affected son. The son of a carrier mother has a 50% chance of inheriting the defective gene from his mother. The daughter of a carrier mother has a 50% chance of being a carrier and a 50% chance of having two normal copies of the gene.

30 In all cases, an unaffected father will either pass a normal Y to his son or a normal X to his

daughter. Female carriers of an X-linked recessive condition, such as DMD, can show symptoms depending on their pattern of X-inactivation.

Exon deletions preceding exon 51 of the human *DMD* gene, which disrupt the open reading frame (ORF) by juxtaposing out of frame exons, represent the most common type of human DMD mutation. Skipping of exon 51 can, in principle, restore the DMD ORF in 13% of DMD patients with exon deletions.

Duchenne muscular dystrophy has an incidence of 1 in 3,500 male infants. Mutations within the dystrophin gene can either be inherited or occur spontaneously during germline transmission.

D. Diagnosis

Genetic counseling is advised for people with a family history of the disorder. Duchenne muscular dystrophy can be detected with about 95% accuracy by genetic studies performed during pregnancy.

DNA test. The muscle-specific isoform of the dystrophin gene is composed of 79 exons, and DNA testing and analysis can usually identify the specific type of mutation of the exon or exons that are affected. DNA testing confirms the diagnosis in most cases.

Muscle biopsy. If DNA testing fails to find the mutation, a muscle biopsy test may be performed. A small sample of muscle tissue is extracted (usually with a scalpel instead of a needle) and a dye is applied that reveals the presence of dystrophin. Complete absence of the protein indicates the condition.

Over the past several years DNA tests have been developed that detect more of the many mutations that cause the condition, and muscle biopsy is not required as often to confirm the presence of Duchenne's.

Prenatal tests. DMD is carried by an X-linked recessive gene. Males have only one X chromosome, so one copy of the mutated gene will cause DMD. Fathers cannot pass X-linked traits on to their sons, so the mutation is transmitted by the mother.

If the mother is a carrier, and therefore one of her two X chromosomes has a DMD mutation, there is a 50% chance that a female child will inherit that mutation as one of her two X chromosomes, and be a carrier. There is a 50% chance that a male child will inherit that mutation as his one X chromosome, and therefore have DMD.

Prenatal tests can tell whether their unborn child has the most common mutations. There are many mutations responsible for DMD, and some have not been identified, so

genetic testing only works when family members with DMD have a mutation that has been identified.

Prior to invasive testing, determination of the fetal sex is important; while males are sometimes affected by this X-linked disease, female DMD is extremely rare. This can be achieved by ultrasound scan at 16 weeks or more recently by free fetal DNA testing. Chorion villus sampling (CVS) can be done at 11–14 weeks, and has a 1% risk of miscarriage. Amniocentesis can be done after 15 weeks, and has a 0.5% risk of miscarriage. Fetal blood sampling can be done at about 18 weeks. Another option in the case of unclear genetic test results is fetal muscle biopsy.

E. Treatment

There is no current cure for DMD, and an ongoing medical need has been recognized by regulatory authorities. Phase 1-2a trials with exon skipping treatment for certain mutations have halted decline and produced small clinical improvements in walking. Treatment is generally aimed at controlling the onset of symptoms to maximize the quality of life, and include the following:

- Corticosteroids such as prednisolone and deflazacort increase energy and strength and defer severity of some symptoms.
- Randomized control trials have shown that beta-2-agonists increase muscle strength but do not modify disease progression. Follow-up time for most RCTs on beta2-agonists is only around 12 months and hence results cannot be extrapolated beyond that time frame.
- Mild, non-jarring physical activity such as swimming is encouraged. Inactivity (such as bed rest) can worsen the muscle disease.
- Physical therapy is helpful to maintain muscle strength, flexibility, and function.
- Orthopedic appliances (such as braces and wheelchairs) may improve mobility and the ability for self-care. Form-fitting removable leg braces that hold the ankle in place during sleep can defer the onset of contractures.
- Appropriate respiratory support as the disease progresses is important.

Comprehensive multi-disciplinary care standards/guidelines for DMD have been developed by the Centers for Disease Control and Prevention (CDC), and are available at www.treat-nmd.eu/dmd/care/diagnosis-management-DMD.

DMD generally progresses through five stages, as outlined in Bushby *et al.*, Lancet Neurol., 9(1): 77-93 (2010) and Bushby *et al.*, Lancet Neurol., 9(2): 177-198 (2010), incorporated by reference in their entireties. During the presymptomatic stage, patients typically show developmental delay, but no gait disturbance. During the early ambulatory stage, patients typically show the Gowers' sign, waddling gait, and toe walking. During the late ambulatory stage, patients typically exhibit an increasingly labored gait and begin to lose the ability to climb stairs and rise from the floor. During the early non-ambulatory stage, patients are typically able to self-propel for some time, are able to maintain posture, and may develop scoliosis. During the late non-ambulatory stage, upper limb function and postural maintenance is increasingly limited.

In some embodiments, treatment is initiated in the presymptomatic stage of the disease. In some embodiments, treatment is initiated in the early ambulatory stage. In some embodiments, treatment is initiated in the late ambulatory stage. In embodiments, treatment is initiated during the early non-ambulatory stage. In embodiments, treatment is initiated during the late non-ambulatory stage.

1. Physical Therapy

Physical therapists are concerned with enabling patients to reach their maximum physical potential. Their aim is to:

- minimize the development of contractures and deformity by developing a program of stretches and exercises where appropriate
- anticipate and minimize other secondary complications of a physical nature by recommending bracing and durable medical equipment
- monitor respiratory function and advise on techniques to assist with breathing exercises and methods of clearing secretions

2. Respiration Assistance

Modern "volume ventilators/respirators," which deliver an adjustable volume (amount) of air to the person with each breath, are valuable in the treatment of people with muscular dystrophy related respiratory problems. The ventilator may require an invasive endotracheal or tracheotomy tube through which air is directly delivered, but, for some people non-invasive delivery through a face mask or mouthpiece is sufficient. Positive airway pressure machines, particularly bi-level ones, are sometimes used in this latter way. The

respiratory equipment may easily fit on a ventilator tray on the bottom or back of a power wheelchair with an external battery for portability.

Ventilator treatment may start in the mid to late teens when the respiratory muscles can begin to collapse. If the vital capacity has dropped below 40% of normal, a volume ventilator/respirator may be used during sleeping hours, a time when the person is most likely to be under ventilating (“hypoventilating”). Hypoventilation during sleep is determined by a thorough history of sleep disorder with an oximetry study and a capillary blood gas (See Pulmonary Function Testing). A cough assist device can help with excess mucus in lungs by hyperinflation of the lungs with positive air pressure, then negative pressure to get the mucus up. If the vital capacity continues to decline to less than 30 percent of normal, a volume ventilator/respirator may also be needed during the day for more assistance. The person gradually will increase the amount of time using the ventilator/respirator during the day as needed.

F. Prognosis

Duchenne muscular dystrophy is a progressive disease which eventually affects all voluntary muscles and involves the heart and breathing muscles in later stages. The life expectancy is currently estimated to be around 25, but this varies from patient to patient. Recent advancements in medicine are extending the lives of those afflicted. The *Muscular Dystrophy Campaign*, which is a leading UK charity focusing on all muscle disease, states that “with high standards of medical care young men with Duchenne muscular dystrophy are often living well into their 30s.”

In rare cases, persons with DMD have been seen to survive into the forties or early fifties, with the use of proper positioning in wheelchairs and beds, ventilator support (via tracheostomy or mouthpiece), airway clearance, and heart medications, if required. Early planning of the required supports for later-life care has shown greater longevity in people living with DMD.

Curiously, in the mdx mouse model of Duchenne muscular dystrophy, the lack of dystrophin is associated with increased calcium levels and skeletal muscle myonecrosis. The intrinsic laryngeal muscles (ILM) are protected and do not undergo myonecrosis. ILM have a calcium regulation system profile suggestive of a better ability to handle calcium changes in comparison to other muscles, and this may provide a mechanistic insight for their unique pathophysiological properties. The ILM may facilitate the development of novel strategies for the prevention and treatment of muscle wasting in a variety of clinical scenarios.

II. CRISPR Systems

A. CRISPRs

CRISPRs (clustered regularly interspaced short palindromic repeats) are DNA loci
5 containing short repetitions of base sequences. Each repetition is followed by short segments
of “spacer DNA” from previous exposures to a virus. CRISPRs are found in approximately
40% of sequenced eubacteria genomes and 90% of sequenced archaea. CRISPRs are often
associated with *cas* genes that code for proteins related to CRISPRs. The CRISPR/Cas system
10 is a prokaryotic immune system that confers resistance to foreign genetic elements such as
plasmids and phages and provides a form of acquired immunity. CRISPR spacers recognize
and silence these exogenous genetic elements like RNAi in eukaryotic organisms.

CRISPR repeats range in size from 24 to 48 base pairs. They usually show some dyad
symmetry, implying the formation of a secondary structure such as a hairpin, but are not truly
palindromic. Repeats are separated by spacers of similar length. Some CRISPR spacer
15 sequences exactly match sequences from plasmids and phages, although some spacers match
the prokaryote’s genome (self-targeting spacers). New spacers can be added rapidly in
response to phage infection.

B. Cas Nucleases

CRISPR-associated (*cas*) genes are often associated with CRISPR repeat-spacer
20 arrays. As of 2013, more than forty different Cas protein families had been described. Of
these protein families, Cas1 appears to be ubiquitous among different CRISPR/Cas systems.
Particular combinations of *cas* genes and repeat structures have been used to define 8
CRISPR subtypes (*Ecoli*, *Ypest*, *Nmeni*, *Dvulg*, *Tneap*, *Hmari*, *Apern*, and *Mtube*), some of
which are associated with an additional gene module encoding repeat-associated mysterious
25 proteins (RAMPs). More than one CRISPR subtype may occur in a single genome. The
sporadic distribution of the CRISPR/Cas subtypes suggests that the system is subject to
horizontal gene transfer during microbial evolution.

Exogenous DNA is apparently processed by proteins encoded by *Cas* genes into small
elements (~30 base pairs in length), which are then somehow inserted into the CRISPR locus
30 near the leader sequence. RNAs from the CRISPR loci are constitutively expressed and are
processed by *Cas* proteins to small RNAs composed of individual, exogenously-derived
sequence elements with a flanking repeat sequence. The RNAs guide other *Cas* proteins to
silence exogenous genetic elements at the RNA or DNA level. Evidence suggests functional
diversity among CRISPR subtypes. The *Cse* (*Cas* subtype *Ecoli*) proteins (called *CasA-E* in

E. coli) form a functional complex, Cascade, that processes CRISPR RNA transcripts into spacer-repeat units that Cascade retains. In other prokaryotes, Cas6 processes the CRISPR transcripts. Interestingly, CRISPR-based phage inactivation in *E. coli* requires Cascade and Cas3, but not Cas1 and Cas2. The Cmr (Cas RAMP module) proteins found in *Pyrococcus*
5 *furiosus* and other prokaryotes form a functional complex with small CRISPR RNAs that recognizes and cleaves complementary target RNAs. RNA-guided CRISPR enzymes are classified as type V restriction enzymes.

Cas9 is a nuclease, an enzyme specialized for cutting DNA, with two active cutting sites, one for each strand of the double helix. The team demonstrated that they could disable
10 one or both sites while preserving Cas9's ability to locate its target DNA. combined tracrRNA and spacer RNA into a "single-guide RNA" molecule that, mixed with Cas9, could find and cut the correct DNA targets proposed that such synthetic guide RNAs might be able to be used for gene editing.

Cas9 proteins are highly enriched in pathogenic and commensal bacteria.
15 CRISPR/Cas-mediated gene regulation may contribute to the regulation of endogenous bacterial genes, particularly during bacterial interaction with eukaryotic hosts. For example, Cas protein Cas9 of *Francisella novicida* uses a unique, small, CRISPR/Cas-associated RNA (scaRNA) to repress an endogenous transcript encoding a bacterial lipoprotein that is critical for *F. novicida* to dampen host response and promote virulence. Wang *et al.* showed that
20 coinjection of Cas9 mRNA and sgRNAs into the germline (zygotes) generated mice with mutations. Delivery of Cas9 DNA sequences also is contemplated.

The systems CRISPR/Cas are separated into three classes. Class I uses several Cas proteins together with the CRISPR RNAs (crRNA) to build a functional endonuclease. Class 2 CRISPR systems use a single Cas protein with a crRNA. Cpf1 has been recently identified
25 as a Class II, Type V CRISPR/Cas systems containing a 1,300 amino acid protein. See also U.S. Patent Publication 2014/0068797, which is incorporated by reference in its entirety.

C. Cpf1 Nucleases

Clustered Regularly Interspaced Short Palindromic Repeats from *Prevotella* and *Francisella* 1 or CRISPR/Cpf1 is a DNA-editing technology which shares some similarities
30 with the CRISPR/Cas9 system. Cpf1 is an RNA-guided endonuclease of a class II CRISPR/Cas system. This acquired immune mechanism is found in *Prevotella* and *Francisella* bacteria. It prevents genetic damage from viruses. Cpf1 genes are associated with the CRISPR locus, coding for an endonuclease that use a guide RNA to find and cleave viral

DNA. Cpf1 is a smaller and simpler endonuclease than Cas9, overcoming some of the CRISPR/Cas9 system limitations.

Cpf1 appears in many bacterial species. The ultimate Cpf1 endonuclease that was developed into a tool for genome editing was taken from one of the first 16 species known to harbor it.

In embodiments, the Cpf1 is a Cpf1 enzyme from *Acidaminococcus* (species BV3L6, UniProt Accession No. U2UMQ6; SEQ ID NO. 442), having the sequence set forth below:

```

1  mtqfegftnl yqvsktlrfe lipqgktrlkh iqeggfieed karndhykel kpiidriykt
61  yadqclqlvq ldwenlsaai dsyrkektee trnalieeqa tyrnaihydyf igrtdnlttda
10  121 inkrhaeiya glfkaelfng kvlkqlgtvt ttehenallr sdfkfttyfs gfyenrknvf
181 saedistaip hrivqdnfnpk fkenchiftr litavpslre hfenvkkaig ifvstsieev
241 fsfpfynqll tqtdidlynq llggisreag tekikglnev lnlaigknde tahiiaslph
301 rfipflfkqil sdrntlsfil eefksdeevi qsfckykttl rnenvletae alfnelnsid
361 lthifishkk letissalcd hwdtlrnaly erriseltgk itksakekvq rslkhedinl
15  421 qeiiisaagke lseafkqkts eilshahaal dqplpttlkk qeekeilksq ldsllglyhl
481 ldwfvavdesn evdpefsarl tgiklemeps lsfnkarny atkpkysvek fklmfqmpnl
541 asgwdvnkek nngailfvkn glyylgimpk qkgrykalsf eptektsegef dkmyydyfpe
601 aakmipkcst qlkavtahfq thtppillsn nfiepleitk eiylndnpek epkkfqtaya
661 kktgdqkgyr ealckwidft rdflskytkt tsidlsslrp ssqykdlgey yaelpnlllyh
20  721 isfqriaeke imdavetgkl ylfqiynkdf akghhgkpnl htlywtglfs penlaktsik
781 lnggaelfyr pksrmkrmah rlgekmnlkk lkdqktpipd tlyqelydyv nhrslhdsld
841 earallpnvi tkevsheiiik drrftsdkff fhvptitlny aanspskfng rvnaylkehpe
901 etpiigidrg ernliytivi dstgkileqr slntiqqfdy qkkldnreke rvaarqawsv
961 vgtikdlkqg ylsqviheiv dlmihyqavv vlenlnfgfk skrtgiaeka vyqqfekmli
25  1021 dklnclvlkd ypaekvggvl npyqltdqft sfakmgtqsg flfyvpapyt skidpltgfv
1081 dpfvwktikn hesrkhfleg fdfilhydvkt gdfilhfkmn rnlsfqrqlp gfm pawdivf
1141 eknetqfdak gtpfiagkri vpvienhrft gryrdlypan elialleekg ivfrdgsnil
1201 pkllenddsh aidtmvalir svlqmrnsna atgedyinsp vrdlngvcfd srfqnpewpm
1261 dadangayhi alkgqlllnh lkeskdiklq ngisnqdwla yiqelnrn
30

```

In some embodiments, the Cpf1 is a Cpf1 enzyme from *Lachnospiraceae* (species ND2006, UniProt Accession No. A0A182DWE3; SEQ ID NO. 443), having the sequence set forth below:

```

1  AASKLEKFTN CYSLSKTLRF KAIPVGKTOE NIDNKRLLE DEKRAEDYKG VKKLLDRYYL
35  61 SFINDVLHSI KLKLNLYNIS LFRKKTRTEK ENKELENLEI NLRKEIAKAF KGAAGYKSLF
121 KKDIETILP EAADDKDEIA LVNSFNGFTT AFTGFFDNRE NMFSEEAKST SIAFRCINEN
181 LTRYISNMDI FEKVDAIFDK HEVQEIKEKI LNSDYDVEDF FEGEFFNFVL TQEGIDVYNA
241 IIGGFVTESE EKIKGLNEYI NLYNAKTQA LPKFKPLYKQ VLSDRESLSF YGEGYTSDEE
301 VLEVFRNTLN KNSEIFSSIK KLEKLFKNFD EYSSAGIFVK NGPAISTISK DIFGEWNLIR
40  361 DKWNAEYDDI HLKKKAVVTE KYEDDRRKSF KKIGSFSLEQ LQYADADLS VVEKLKEIII
421 QKVDEIYKVY GSSEKLFAD FVLEKSLKKN DAVVAIMKDL LDSVKSFENY IKAFFGEGKE
481 TNRDESFYGD FVLAYDILLK VDHIYDAIRN YVTQKPYSD KFKLYFQNPQ FMGGWDKDKKE
541 TDYRATILRY GSKYYLAIMD KKYAKCLQKI DKDDVNGNYE KINYKLLPGP NKMLPKVFFS
601 KKWMAYNPS EDIQKIYKNG TFKKGDMFNL NDCHKLIDFF KDSISRYPKW SNAYDFNFSE
45  661 TEKYKDIAGF YREVEEQGYK VSFESASKKE VDKLVEEGKL YMFQIYNKDF SDKSHGTPNL
721 HTMYFKLLFD ENNHGQIRLS GGAELEMRRA SLKKEELVH PANSPIANKN PDNPKKTTTL
781 SYDVYKDKRF SEDQYELHIP IAINCKPKNI FKINTEVRVL LKHDDNPYVI GIDRGERNLL
841 YIVVVDGKGN IVEQYSLNEI INNFNIRIK TDYHSLDDKK EKERFEARQN WTSIENIKEL
901 KAGYISQVH KICELVEKYD AVIALEDLNS GFKNSRVKVE KQVYQKFEKM LIDKLYNMVD
50  961 KKSNPCATGG ALKGYQITNK FESFKSMSTQ NGFIFYIPAW LSKIDPSTG FVNLLKTKYT
1021 SIADSKFIS SFDRIMYVPE EDLFEALDY KNFSRTDADY IKKWKLYSYG NRIRIFAAAK
1081 KNNVFAWEV CLTSAYKELF NKYGINYQQG DIRALLCEQS DKAIFYSSFMA LMSLMLQMRN
1141 SITGRTDVDF LISPVKNSDG IFYDSRNYEA QENAILPKNA DANGAYNIAR KVLWAIQQFK

```

1201 KAEDEKLDKV KIAISNKEWL EYAQTSVK

In some embodiments, the Cpf1 is codon optimized for expression in mammalian cells. In some embodiments, the Cpf1 is codon optimized for expression in human cells.

5 In some embodiments, the compositions of the disclosure include a small version of a Cas9 from the bacterium *Staphylococcus aureus*. The small version of the Cas9 provides advantages over wild type or full length Cas9.

The Cpf1 locus contains a mixed alpha/beta domain, a RuvC-I followed by a helical region, a RuvC-II and a zinc finger-like domain. The Cpf1 protein has a RuvC-like
10 endonuclease domain that is similar to the RuvC domain of Cas9. Furthermore, Cpf1 does not have a HNH endonuclease domain, and the N-terminal of Cpf1 does not have the alpha-helical recognition lobe of Cas9.

Cpf1 CRISPR-Cas domain architecture shows that Cpf1 is functionally unique, being classified as Class 2, type V CRISPR system. The Cpf1 loci encode Cas1, Cas2 and Cas4
15 proteins more similar to types I and III than from type II systems. Database searches suggest the abundance of Cpf1-family proteins in many bacterial species.

Functional Cpf1 does not require a tracrRNA. Therefore, functional Cpf1 gRNAs of the disclosure may comprise or consist of a crRNA. This benefits genome editing because Cpf1 is not only a smaller nuclease than Cas9, but also it has a smaller sgRNA molecule
20 (proximately half as many nucleotides as Cas9).

The Cpf1-gRNA (e.g. Cpf1-crRNA) complex cleaves target DNA or RNA by identification of a protospacer adjacent motif 5'-YTN-3' (where "Y" is a pyrimidine and "N" is any nucleobase) or 5'-TTN-3', in contrast to the G-rich PAM targeted by Cas9. After identification of PAM, Cpf1 introduces a sticky-end-like DNA double-stranded break of 4 or
25 5 nucleotides overhang.

The CRISPR/Cpf1 system comprises or consists of a Cpf1 enzyme and a guide RNA that finds and positions the complex at the correct spot on the double helix to cleave target DNA. In its native bacterial hosts, CRISPR/Cpf1 systems activity has three stages:

Adaptation, during which Cas1 and Cas2 proteins facilitate the adaptation of small
30 fragments of DNA into the CRISPR array;

Formation of crRNAs: processing of pre-cr-RNAs producing of mature crRNAs to guide the Cas protein; and

Interference, in which the Cpf1 is bound to a crRNA to form a binary complex to identify and cleave a target DNA sequence.

This system has been modified to utilize non-naturally occurring crRNAs, which guide Cpf1 to a desired target sequence in a non-bacterial cell, such as a mammalian cell.

D. gRNA

5 As an RNA guided protein, Cas9 requires a short RNA to direct the recognition of DNA targets. Though Cas9 preferentially interrogates DNA sequences containing a PAM sequence NGG it can bind here without a protospacer target. However, the Cas9-gRNA complex requires a close match to the gRNA to create a double strand break. CRISPR sequences in bacteria are expressed in multiple RNAs and then processed to create guide
10 strands for RNA. Because Eukaryotic systems lack some of the proteins required to process CRISPR RNAs the synthetic construct gRNA was created to combine the essential pieces of RNA for Cas9 targeting into a single RNA expressed with the RNA polymerase type III promoter U6. Synthetic gRNAs are slightly over 100bp at the minimum length and contain a portion which is targets the 20 protospacer nucleotides immediately preceding the PAM
15 sequence NGG; gRNAs do not contain a PAM sequence.

In some embodiments, the gRNA targets a site within a wildtype dystrophin gene. In some embodiments, the gRNA targets a site within a mutant dystrophin gene. In some embodiments, the gRNA targets a dystrophin intron. In some embodiments, the gRNA targets dystrophin exon. In some embodiments, the gRNA targets a site in a dystrophin exon
20 that is expressed and is present in one or more of the dystrophin isoforms shown in Table 1. In embodiments, the gRNA targets a dystrophin splice site. In some embodiments, the gRNA targets a splice donor site on the dystrophin gene. In embodiments, the gRNA targets a splice acceptor site on the dystrophin gene.

In embodiments, the guide RNA targets a mutant DMD exon. In some embodiments,
25 the mutant exon is exon 23 or 51. In some embodiments, the guide RNA targets at least one of exons 1, 23, 41, 44, 46, 47, 48, 49, 50, 51, 52, 53, 54, or 55 of the dystrophin gene. In embodiments, the guide RNA targets at least one of introns 44, 45, 50, 51, 52, 53, 54, or 55 of the dystrophin gene. In preferred embodiments, the guide RNAs are designed to induce skipping of exon 51 or exon 23. In embodiments, the gRNA is targeted to a splice acceptor
30 site of exon 51 or exon 23.

Suitable gRNAs for use in the methods and compositions disclosed herein are provided as SEQ ID NOs. 60-382. (Table E). In preferred embodiments, the gRNA is selected from any one of SEQ ID No. 60 to SEQ ID No. 382.

In some embodiments, gRNAs of the disclosure comprise a sequence that is complementary to a target sequence within a coding sequence or a non-coding sequence corresponding to the *DMD* gene, and, therefore, hybridize to the target sequence. In some embodiments, gRNAs for Cpf1 comprise a single crRNA containing a direct repeat scaffold sequence followed by 24 nucleotides of guide sequence. In some embodiments, a “guide” sequence of the crRNA comprises a sequence of the gRNA that is complementary to a target sequence. In some embodiments, crRNA of the disclosure comprises a sequence of the gRNA that is not complementary to a target sequence. “Scaffold” sequences of the disclosure link the gRNA to the Cpf1 polypeptide. “Scaffold” sequences of the disclosure are not equivalent to a tracrRNA sequence of a gRNA-Cas9 construct.

E. Cas9 versus Cpf1

Cas9 requires two RNA molecules to cut DNA while Cpf1 needs one. The proteins also cut DNA at different places, offering researchers more options when selecting an editing site. Cas9 cuts both strands in a DNA molecule at the same position, leaving behind ‘blunt’ ends. Cpf1 leaves one strand longer than the other, creating ‘sticky’ ends that are easier to work with. Cpf1 appears to be more able to insert new sequences at the cut site, compared to Cas9. Although the CRISPR/Cas9 system can efficiently disable genes, it is challenging to insert genes or generate a knock-in. Cpf1 lacks tracrRNA, utilizes a T-rich PAM and cleaves DNA via a staggered DNA DSB.

In summary, important differences between Cpf1 and Cas9 systems are that Cpf1 recognizes different PAMs, enabling new targeting possibilities, creates 4-5 nt long sticky ends, instead of blunt ends produced by Cas9, enhancing the efficiency of genetic insertions and specificity during NHEJ or HDR, and cuts target DNA further away from PAM, further away from the Cas9 cutting site, enabling new possibilities for cleaving the DNA.

<u>Feature</u>	<u>Cas9</u>	<u>Cpf1</u>
Structure	Two RNA required (Or 1 fusion transcript (crRNA+tracrRNA=gRNA))	One RNA required
Cutting mechanism	Blunt end cuts	Staggered end cuts
Cutting site	Proximal to recognition site	Distal from recognition site
Target sites	G-rich PAM	T-rich PAM
Cell type	Fast growing cells, including cancer cells	Non-dividing cells, including nerve cells

F. CRISPR/Cpf1-mediated gene editing

The first step in editing the DMD gene using CRISPR/Cpf1 is to identify the genomic target sequence. The genomic target for the gRNAs of the disclosure can be any ~24 nucleotide DNA sequence within the dystrophin gene, provided that the sequence is unique compared to the rest of the genome. In some embodiments, the genomic target sequence is in exon 51, exon 45, exon 44, exon 53, exon 46, exon 52, exon 50, exon 43, exon 6, exon 7, exon 8, and/or exon 55 of the human dystrophin gene. In some embodiments, the genomic target sequence is a 5' or 3' splice site of exon 51, exon 45, exon 44, exon 53, exon 46, exon 52, exon 50, exon 43, exon 6, exon 7, exon 8, and/or exon 55 of the human dystrophin gene. In some embodiments, the genomic target sequence is an intron immediately upstream or downstream of exon 51, exon 45, exon 44, exon 53, exon 46, exon 52, exon 50, exon 43, exon 6, exon 7, exon 8, and/or exon 55 of the human dystrophin gene. Exemplary genomic target sequences can be found in Table D.

The next step in editing the DMD gene using CRISPR/Cpf1 is to identify all Protospacer Adjacent Motif (PAM) sequences within the genetic region to be targeted. Cpf1 utilizes a T-rich PAM sequence (TTTN, wherein N is any nucleotide). The target sequence must be immediately upstream of a PAM. Once all possible PAM sequences and putative target sites have been identified, the next step is to choose which site is likely to result in the most efficient on-target cleavage. The gRNA targeting sequence needs to match the target sequence, and the gRNA targeting sequence must not match additional sites within the genome. In preferred embodiments, the gRNA targeting sequence has perfect homology to the target with no homology elsewhere in the genome. In some embodiments, a given gRNA targeting sequence will have additional sites throughout the genome where partial homology exists. These sites are called "off-targets" and should be considered when designing a gRNA. In general, off-target sites are not cleaved as efficiently when mismatches occur near the PAM sequence, so gRNAs with no homology or those with mismatches close to the PAM sequence will have the highest specificity. In addition to "off-target activity", factors that maximize cleavage of the desired target sequence ("on-target activity") must be considered. It is known to those of skill in the art that two gRNA targeting sequences, each having 100% homology to the target DNA may not result in equivalent cleavage efficiency. In fact, cleavage efficiency may increase or decrease depending upon the specific nucleotides within the selected target sequence. Close examination of predicted on-target and off-target activity of each potential gRNA targeting sequence is necessary to design the best gRNA. Several gRNA design programs have been developed that are capable of locating potential PAM and

target sequences and ranking the associated gRNAs based on their predicted on-target and off-target activity (e.g. CRISPRdirect, available at www.crispr.dbcls.jp).

The next step is to synthesize and clone desired gRNAs. Targeting oligos can be synthesized, annealed, and inserted into plasmids containing the gRNA scaffold using standard restriction-ligation cloning. However, the exact cloning strategy will depend on the gRNA vector that is chosen. The gRNAs for Cpf1 are notably simpler than the gRNAs for Cas9, and only consist of a single crRNA containing direct repeat scaffold sequence followed by ~24 nucleotides of guide sequence. Cpf1 requires a minimum of 16 nucleotides of guide sequence to achieve detectable DNA cleavage, and a minimum of 18 nucleotides of guide sequence to achieve efficient DNA cleavage in vitro. In some embodiments, 20-24 nucleotides of guide sequence is used. The seed region of the Cpf1 gRNA is generally within the first 5 nucleotides on the 5' end of the guide sequence. Cpf1 makes a staggered cut in the target genomic DNA. In AsCpf1 and LbCpf1, the cut occurs 19 bp after the PAM on the targeted (+) strand, and 23 bp on the other strand.

Each gRNA should then be validated in one or more target cell lines. For example, after the CRISPR and gRNA are delivered to the cell, the genomic target region may be amplified using PCR and sequenced according to methods known to those of skill in the art.

In some embodiments, gene editing may be performed in vitro or ex vivo. In some embodiments, cells are contacted in vitro or ex vivo with a Cpf1 and a gRNA that targets a dystrophin splice site. In some embodiments, the cells are contacted with one or more nucleic acids encoding the Cpf1 and the guide RNA. In some embodiments, the one or more nucleic acids are introduced into the cells using, for example, lipofection or electroporation. Gene editing may also be performed in zygotes. In embodiments, zygotes may be injected with one or more nucleic acids encoding Cpf1 and a gRNA that targets a dystrophin splice site. The zygotes may subsequently be injected into a host.

In embodiments, the Cpf1 is provided on a vector. In embodiments, the vector contains a Cpf1 sequence derived from a *Lachnospiraceae* bacterium. See, for example, Uniprot Accession No. A0A182DWE3; SEQ ID NO. 443. In embodiments, the vector contains a Cpf1 sequence derived from an *Acidaminococcus* bacterium. See, for example, Uniprot Accession No. U2UMQ6; SEQ ID NO. 442. In some embodiments, the Cpf1 sequence is codon optimized for expression in human cells. In some embodiments, the vector further contains a sequence encoding a fluorescent protein, such as GFP, which allows Cpf1-expressing cells to be sorted using fluorescence activated cell sorting (FACS). In some embodiments, the vector is a viral vector such as an adeno-associated viral vector.

In embodiments, the gRNA is provided on a vector. In some embodiments, the vector is a viral vector such as an adeno-associated viral vector. In embodiments, the Cpf1 and the guide RNA are provided on the same vector. In embodiments, the Cpf1 and the guide RNA are provided on different vectors.

5 In some embodiments, the cells are additionally contacted with a single-stranded DMD oligonucleotide to effect homology directed repair. In some embodiments, small INDELs restore the protein reading frame of dystrophin ("reframing" strategy). When the reframing strategy is used, the cells may be contacted with a single gRNA. In embodiments, a splice donor or splice acceptor site is disrupted, which results in exon skipping and
10 restoration of the protein reading frame ("exon skipping" strategy). When the exon skipping strategy is used, the cells may be contacted with two or more gRNAs.

Efficiency of in vitro or ex vivo Cpf1-mediated DNA cleavage may be assessed using techniques known to those of skill in the art, such as the T7 E1 assay. Restoration of DMD expression may be confirmed using techniques known to those of skill in the art, such as RT-
15 PCR, western blotting, and immunocytochemistry.

In some embodiments, in vitro or ex vivo gene editing is performed in a muscle or satellite cell. In some embodiments, gene editing is performed in iPSC or iCM cells. In embodiments, the iPSC cells are differentiated after gene editing. For example, the iPSC cells may be differentiated into a muscle cell or a satellite cell after editing. In embodiments,
20 the iPSC cells are differentiated into cardiac muscle cells, skeletal muscle cells, or smooth muscle cells. In embodiments, the iPSC cells are differentiated into cardiomyocytes. iPSC cells may be induced to differentiate according to methods known to those of skill in the art.

In some embodiments, contacting the cell with the Cpf1 and the gRNA restores dystrophin expression. In embodiments, cells which have been edited in vitro or ex vivo, or
25 cells derived therefrom, show levels of dystrophin protein that is comparable to wild type cells. In embodiments, the edited cells, or cells derived therefrom, express dystrophin at a level that is 50%, 60%, 70%, 80%, 90%, 95% or any percentage in between of wild type dystrophin expression levels. In embodiments, the cells which have been edited in vitro or ex vivo, or cells derived therefrom, have a mitochondrial number that is comparable to that of
30 wild type cells. In embodiments the edited cells, or cells derived therefrom, have 50%, 60%, 70%, 80%, 90%, 95% or any percentage in between as many mitochondria as wild type cells. In embodiments, the edited cells, or cells derived therefrom, show an increase in oxygen consumption rate (OCR) compared to non-edited cells at baseline.

III. Nucleic Acid Delivery

As discussed above, in certain embodiments, expression cassettes are employed to express a transcription factor product, either for subsequent purification and delivery to a cell/subject, or for use directly in a genetic-based delivery approach. Provided herein are expression vectors which contain one or more nucleic acids encoding Cpf1 and at least one DMD guide RNA that targets a dystrophin splice site. In some embodiments, a nucleic acid encoding Cpf1 and a nucleic acid encoding at least one guide RNA are provided on the same vector. In further embodiments, a nucleic acid encoding Cpf1 and a nucleic acid encoding least one guide RNA are provided on separate vectors.

Expression requires that appropriate signals be provided in the vectors, and include various regulatory elements such as enhancers/promoters from both viral and mammalian sources that drive expression of the genes of interest in cells. Elements designed to optimize messenger RNA stability and translatability in host cells also are defined. The conditions for the use of a number of dominant drug selection markers for establishing permanent, stable cell clones expressing the products are also provided, as is an element that links expression of the drug selection markers to expression of the polypeptide.

A. Regulatory Elements

Throughout this application, the term “expression cassette” is meant to include any type of genetic construct containing a nucleic acid coding for a gene product in which part or all of the nucleic acid encoding sequence is capable of being transcribed and translated, *i.e.*, is under the control of a promoter. A “promoter” refers to a DNA sequence recognized by the synthetic machinery of the cell, or introduced synthetic machinery, required to initiate the specific transcription of a gene. The phrase “under transcriptional control” means that the promoter is in the correct location and orientation in relation to the nucleic acid to control RNA polymerase initiation and expression of the gene. An “expression vector” is meant to include expression cassettes comprised in a genetic construct that is capable of replication, and thus including one or more of origins of replication, transcription termination signals, poly-A regions, selectable markers, and multipurpose cloning sites.

The term promoter will be used here to refer to a group of transcriptional control modules that are clustered around the initiation site for RNA polymerase II. Much of the thinking about how promoters are organized derives from analyses of several viral promoters, including those for the HSV thymidine kinase (*tk*) and SV40 early transcription units. These studies, augmented by more recent work, have shown that promoters are composed of

discrete functional modules, each consisting of approximately 7-20 bp of DNA, and containing one or more recognition sites for transcriptional activator or repressor proteins.

At least one module in each promoter functions to position the start site for RNA synthesis. The best known example of this is the TATA box, but in some promoters lacking
5 a TATA box, such as the promoter for the mammalian terminal deoxynucleotidyl transferase gene and the promoter for the SV40 late genes, a discrete element overlying the start site itself helps to fix the place of initiation.

In some embodiments, the Cpf1 constructs of the disclosure are expressed by a muscle-cell specific promoter. This muscle-cell specific promoter may be constitutively
10 active or may be an inducible promoter.

Additional promoter elements regulate the frequency of transcriptional initiation. Typically, these are located in the region 30-110 bp upstream of the start site, although a number of promoters have recently been shown to contain functional elements downstream of the start site as well. The spacing between promoter elements frequently is flexible, so that
15 promoter function is preserved when elements are inverted or moved relative to one another. In the tk promoter, the spacing between promoter elements can be increased to 50 bp apart before activity begins to decline. Depending on the promoter, it appears that individual elements can function either co-operatively or independently to activate transcription.

In certain embodiments, viral promoters such as the human cytomegalovirus (CMV)
20 immediate early gene promoter, the SV40 early promoter, the Rous sarcoma virus long terminal repeat, rat insulin promoter and glyceraldehyde-3-phosphate dehydrogenase can be used to obtain high-level expression of the coding sequence of interest. The use of other viral or mammalian cellular or bacterial phage promoters which are well-known in the art to achieve expression of a coding sequence of interest is contemplated as well, provided that the
25 levels of expression are sufficient for a given purpose. By employing a promoter with well-known properties, the level and pattern of expression of the protein of interest following transfection or transformation can be optimized. Further, selection of a promoter that is regulated in response to specific physiologic signals can permit inducible expression of the gene product.

30 Enhancers are genetic elements that increase transcription from a promoter located at a distant position on the same molecule of DNA. Enhancers are organized much like promoters. That is, they are composed of many individual elements, each of which binds to one or more transcriptional proteins. The basic distinction between enhancers and promoters is operational. An enhancer region as a whole must be able to stimulate transcription at a

distance; this need not be true of a promoter region or its component elements. On the other hand, a promoter must have one or more elements that direct initiation of RNA synthesis at a particular site and in a particular orientation, whereas enhancers lack these specificities.

Promoters and enhancers are often overlapping and contiguous, often seeming to have a very similar modular organization.

The promoter and/or enhancer may be, for example, immunoglobulin light chain, immunoglobulin heavy chain, T-cell receptor, HLA DQ α and/or DQ β , β -interferon, interleukin-2, interleukin-2 receptor, MHC class II β , MHC class II HLA-Dra, β -Actin, muscle creatine kinase (MCK), prealbumin (transthyretin), elastase I, metallothionein (MTII), collagenase, albumin, α -fetoprotein, t-globin, β -globin, c-fos, c-HA-*ras*, insulin, neural cell adhesion molecule (NCAM), α_1 -antitrypsin, H2B (TH2B) histone, mouse and/or type I collagen, glucose-regulated proteins (GRP94 and GRP78), rat growth hormone, human serum amyloid A (SAA), troponin I (TN I), platelet-derived growth factor (PDGF), duchenne muscular dystrophy, SV40, polyoma, retroviruses, papilloma virus, hepatitis B virus, human immunodeficiency virus, cytomegalovirus (CMV), and gibbon ape leukemia virus.

In some embodiments, inducible elements may be used. In some embodiments, the inducible element is, for example, MTII, MMTV (mouse mammary tumor virus), β -interferon, adenovirus 5 E2, collagenase, stromelysin, SV40, murine MX gene, GRP78 gene, α -2-macroglobulin, vimentin, MHC class I gene H-2kb, HSP70, proliferin, tumor necrosis factor, and/or thyroid stimulating hormone α gene. In some embodiments, the inducer is phorbol ester (TFA), heavy metals, glucocorticoids, poly(rI)x, poly(rc), EIA, phorbol ester (TPA), interferon, Newcastle Disease Virus, A23187, IL-6, serum, interferon, SV40 large T antigen, PMA, and/or thyroid hormone. Any of the inducible elements described herein may be used with any of the inducers described herein.

Below is a list of promoters/enhancers and inducible promoters/enhancers that could be used in combination with the nucleic acid encoding a gene of interest in an expression construct. Additionally, any promoter/enhancer combination (as per the Eukaryotic Promoter Data Base EPDB) could also be used to drive expression of the gene. Eukaryotic cells can support cytoplasmic transcription from certain bacterial promoters if the appropriate bacterial polymerase is provided, either as part of the delivery complex or as an additional genetic expression construct.

Of particular interest are muscle specific promoters. These include the myosin light chain-2 promoter, the α -actin promoter, the troponin I promoter; the $\text{Na}^+/\text{Ca}^{2+}$ exchanger

promoter, the dystrophin promoter, the $\alpha 7$ integrin promoter, the brain natriuretic peptide promoter and the α B-crystallin/small heat shock protein promoter, α -myosin heavy chain promoter, and the ANF promoter.

In some embodiments, the Cpf1-gRNA constructs of the disclosure are expressed by a muscle-cell specific promoter. This muscle-cell specific promoter may be constitutively active or may be an inducible promoter.

Where a cDNA insert is employed, one will typically desire to include a polyadenylation signal to effect proper polyadenylation of the gene transcript. The nature of the polyadenylation signal is not believed to be crucial to the successful practice of the methods disclosed herein, and any such sequence may be employed such as human growth hormone and SV40 polyadenylation signals. Also contemplated as an element of the expression cassette is a terminator. These elements can serve to enhance message levels and to minimize read through from the cassette into other sequences.

B. 2A Peptide

The inventor utilizes the 2A-like self-cleaving domain from the insect virus *Thosea asigna* (TaV 2A peptide; SEQ ID NO. 444; EGRGSLTCTGDVEENPGP). These 2A-like domains have been shown to function across Eukaryotes and cause cleavage of amino acids to occur co-translationally within the 2A-like peptide domain. Therefore, inclusion of TaV 2A peptide allows the expression of multiple proteins from a single mRNA transcript. Importantly, the domain of TaV when tested in eukaryotic systems have shown greater than 99% cleavage activity (Donnelly *et al.*, 2001). Other acceptable 2A-like peptides include, but are not limited to, equine rhinitis A virus (ERAV) 2A peptide (SEQ ID NO. 445; QCTNYALLKLAGDVESNPGP), porcine teschovirus-1 (PTV1) 2A peptide (SEQ ID NO. 446; ATNFSLLKQAGDVEENPGP) and foot and mouth disease virus (FMDV) 2A peptide (SEQ ID No. 447; PVKQLLNFDLLKLAGDVESNPGP) or modified versions thereof.

In some embodiments, the 2A peptide is used to express a reporter and a Cpf1 simultaneously. The reporter may be, for example, GFP.

Other self-cleaving peptides that may be used include, but are not limited to nuclear inclusion protein a (Nia) protease, a P1 protease, a 3C protease, a L protease, a 3C-like protease, or modified versions thereof.

C. Delivery of Expression Vectors

There are a number of ways in which expression vectors may introduced into cells. In certain embodiments, the expression construct comprises a virus or engineered construct derived from a viral genome. The ability of certain viruses to enter cells via receptor-mediated endocytosis, to integrate into host cell genome and express viral genes stably and efficiently have made them attractive candidates for the transfer of foreign genes into mammalian cells. These have a relatively low capacity for foreign DNA sequences and have a restricted host spectrum. Furthermore, their oncogenic potential and cytopathic effects in permissive cells raise safety concerns. They can accommodate only up to 8 kB of foreign genetic material but can be readily introduced in a variety of cell lines and laboratory animals.

One of the preferred methods for *in vivo* delivery involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences sufficient to (a) support packaging of the construct and (b) to express an antisense polynucleotide that has been cloned therein. In this context, expression does not require that the gene product be synthesized.

The expression vector comprises a genetically engineered form of adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kB, linear, double-stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kB. In contrast to retrovirus, the adenoviral infection of host cells does not result in chromosomal integration because adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be linked only to mild disease such as acute respiratory disease in humans.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication, late gene expression and host

cell shut-off. The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess a 5'-
5 tripartite leader (TPL) sequence which makes them preferred mRNA's for translation.

In one system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be generated from this process. Therefore, it is critical to isolate a single clone of virus from an individual plaque and examine its genomic
10 structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins. Since the E3 region is dispensable from the adenovirus genome, the current
15 adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions. In nature, adenovirus can package approximately 105% of the wild-type genome, providing capacity for about 2 extra kb of DNA. Combined with the approximately 5.5 kb of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kb, or about 15% of the total length of the vector.
20 More than 80% of the adenovirus viral genome remains in the vector backbone and is the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete.

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial
25 cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the preferred helper cell line is 293.

Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and
30 propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is

employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

The adenoviruses of the disclosure are replication defective or at least conditionally replication defective. Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the methods disclosed herein. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup C is the preferred starting material in order to obtain the conditional replication-defective adenovirus vector for use in the present disclosure.

As stated above, the typical vector according to the present disclosure is replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical. The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted E3 region in E3 replacement vectors, as described by Karlsson *et al.* (1986), or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*, 10^9 - 10^{12} plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus does not require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

Adenovirus vectors have been used in eukaryotic gene expression and vaccine development. Animal studies suggested that recombinant adenovirus could be used for gene therapy. Studies in administering recombinant adenovirus to different tissues include trachea instillation, muscle injection, peripheral intravenous injections and stereotactic inoculation into the brain.

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription. The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5' and 3' ends of the viral genome. These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome.

In order to construct a retroviral vector, a nucleic acid encoding a gene of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed. When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be packaged into viral particles, which are then secreted into the culture media. The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells.

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical addition of lactose residues to the viral envelope. This modification could permit the specific infection of hepatocytes via sialoglycoprotein receptors.

A different approach to targeting of recombinant retroviruses may be used, in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor are used. The antibodies are coupled via the biotin components by using streptavidin. Using antibodies against major histocompatibility complex class I and class II antigens, it has been demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

There are certain limitations to the use of retrovirus vectors in all aspects of the present disclosure. For example, retrovirus vectors usually integrate into random sites in the cell genome. This can lead to insertional mutagenesis through the interruption of host genes or through the insertion of viral regulatory sequences that can interfere with the function of flanking genes. Another concern with the use of defective retrovirus vectors is the potential appearance of wild-type replication-competent virus in the packaging cells. This can result from recombination events in which the intact-sequence from the recombinant virus inserts upstream from the gag, pol, env sequence integrated in the host cell genome. However, new packaging cell lines are now available that should greatly decrease the likelihood of recombination (see, for example, Markowitz *et al.*, 1988; Hersdorffer *et al.*, 1990).

Other viral vectors may be employed as expression constructs in the present disclosure. Vectors derived from viruses such as vaccinia virus, adeno-associated virus (AAV) and herpesviruses may be employed. They offer several attractive features for various mammalian cells.

In embodiments, the AAV vector is replication-defector or conditionally replication defective. In embodiments, the AAV vector is a recombinant AAV vector. In some embodiments, the AAV vector comprises a sequence isolated or derived from an AAV vector of serotype AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 or any combination thereof. In some embodiments, the AAV vector is not an AAV9 vector.

In some embodiments, a single viral vector is used to deliver a nucleic acid encoding Cpf1 and at least one gRNA to a cell. In some embodiments, Cpf1 is provided to a cell using a first viral vector and at least one gRNA is provided to the cell using a second viral vector. In order to effect expression of sense or antisense gene constructs, the expression construct must be delivered into a cell. The cell may be a muscle cell, a satellite cell, a mesangioblast, a bone marrow derived cell, a stromal cell or a mesenchymal stem cell. In embodiments, the cell is a cardiac muscle cell, a skeletal muscle cell, or a smooth muscle cell. In embodiments, the cell is a cell in the tibialis anterior, quadriceps, soleus, diaphragm or heart. In some embodiments, the cell is an induced pluripotent stem cell (iPSC) or inner cell mass cell (iCM). In further embodiments, the cell is a human iPSC or a human iCM. In some embodiments, human iPSCs or human iCMs of the disclosure may be derived from a cultured stem cell line, an adult stem cell, a placental stem cell, or from another source of adult or embryonic stem cells that does not require the destruction of a human embryo. Delivery to a cell may be accomplished *in vitro*, as in laboratory procedures for transforming cells lines, or

in vivo or *ex vivo*, as in the treatment of certain disease states. One mechanism for delivery is via viral infection where the expression construct is encapsulated in an infectious viral particle.

Several non-viral methods for the transfer of expression constructs into cultured mammalian cells also are contemplated by the present disclosure. These include calcium phosphate precipitation (Graham and Van Der Eb, 1973; Chen and Okayama, 1987; Rippe *et al.*, 1990) DEAE-dextran (Gopal, 1985), electroporation (Tur-Kaspa *et al.*, 1986; Potter *et al.*, 1984), direct microinjection (Harland and Weintraub, 1985), DNA-loaded liposomes (Nicolau and Sene, 1982; Fraley *et al.*, 1979) and lipofectamine-DNA complexes, cell sonication (Fechheimer *et al.*, 1987), gene bombardment using high velocity microprojectiles (Yang *et al.*, 1990), and receptor-mediated transfection (Wu and Wu, 1987; Wu and Wu, 1988). Some of these techniques may be successfully adapted for *in vivo* or *ex vivo* use.

Once the expression construct has been delivered into the cell the nucleic acid encoding the gene of interest may be positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the gene may be stably integrated into the genome of the cell. This integration may be in the cognate location and orientation via homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or “episomes” encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

In yet another embodiment, the expression construct may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell membrane. This is particularly applicable for transfer *in vitro* but it may be applied to *in vivo* use as well. Dubensky *et al.* (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Neshif (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

In still another embodiment for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them. Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

In some embodiments, the expression construct is delivered directly to the liver, skin, and/or muscle tissue of a subject. This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e.*, *ex vivo* treatment. Again, DNA encoding a particular gene may be delivered via this method and still be incorporated by the present disclosure.

In a further embodiment, the expression construct may be entrapped in a liposome. Liposomes are vesicular structures characterized by a phospholipid bilayer membrane and an inner aqueous medium. Multilamellar liposomes have multiple lipid layers separated by aqueous medium. They form spontaneously when phospholipids are suspended in an excess of aqueous solution. The lipid components undergo self-rearrangement before the formation of closed structures and entrap water and dissolved solutes between the lipid bilayers. Also contemplated are lipofectamine-DNA complexes.

Liposome-mediated nucleic acid delivery and expression of foreign DNA *in vitro* has been very successful. Wong *et al.*, (1980) demonstrated the feasibility of liposome-mediated delivery and expression of foreign DNA in cultured chick embryo, HeLa and hepatoma cells. Nicolau *et al.*, (1987) accomplished successful liposome-mediated gene transfer in rats after intravenous injection. A reagent known as Lipofectamine 2000™ is widely used and commercially available.

In certain embodiments, the liposome may be complexed with a hemagglutinating virus (HVJ). This has been shown to facilitate fusion with the cell membrane and promote cell entry of liposome-encapsulated DNA. In other embodiments, the liposome may be complexed or employed in conjunction with nuclear non-histone chromosomal proteins (HMG-1). In yet further embodiments, the liposome may be complexed or employed in conjunction with both HVJ and HMG-1. In that such expression constructs have been successfully employed in transfer and expression of nucleic acid *in vitro* and *in vivo*, then they are applicable for the present disclosure. Where a bacterial promoter is employed in the

DNA construct, it also will be desirable to include within the liposome an appropriate bacterial polymerase.

Other expression constructs which can be employed to deliver a nucleic acid encoding a particular gene into cells are receptor-mediated delivery vehicles. These take advantage of the selective uptake of macromolecules by receptor-mediated endocytosis in almost all eukaryotic cells. Because of the cell type-specific distribution of various receptors, the delivery can be highly specific.

Receptor-mediated gene targeting vehicles generally consist of two components: a cell receptor-specific ligand and a DNA-binding agent. Several ligands have been used for receptor-mediated gene transfer. The most extensively characterized ligands are asialoorosomucoid (ASOR) and transferrin. A synthetic neoglycoprotein, which recognizes the same receptor as ASOR, has been used as a gene delivery vehicle and epidermal growth factor (EGF) has also been used to deliver genes to squamous carcinoma cells.

IV. Pharmaceutical Compositions and Delivery Methods

For clinical applications, pharmaceutical compositions are prepared in a form appropriate for the intended application. Generally, this entails preparing compositions that are essentially free of pyrogens, as well as other impurities that could be harmful to humans or animals.

Appropriate salts and buffers are used to render drugs, proteins or delivery vectors stable and allow for uptake by target cells. Aqueous compositions of the present disclosure comprise an effective amount of the drug, vector or proteins, dissolved or dispersed in a pharmaceutically acceptable carrier or aqueous medium. The phrase “pharmaceutically or pharmacologically acceptable” refers to molecular entities and compositions that do not produce adverse, allergic, or other untoward reactions when administered to an animal or a human. As used herein, “pharmaceutically acceptable carrier” includes solvents, buffers, solutions, dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents and the like acceptable for use in formulating pharmaceuticals, such as pharmaceuticals suitable for administration to humans. The use of such media and agents for pharmaceutically active substances is well known in the art. Any conventional media or agent that is not incompatible with the active ingredients of the present disclosure, its use in therapeutic compositions may be used. Supplementary active ingredients also can be incorporated into the compositions, provided they do not inactivate the vectors or cells of the compositions.

In some embodiments, the active compositions of the present disclosure include classic pharmaceutical preparations. Administration of these compositions according to the present disclosure may be via any common route so long as the target tissue is available via that route, but generally including systemic administration. This includes oral, nasal, or
5 buccal. Alternatively, administration may be by intradermal, subcutaneous, intramuscular, intraperitoneal or intravenous injection, or by direct injection into muscle tissue. Such compositions are normally administered as pharmaceutically acceptable compositions, as described *supra*.

The active compounds may also be administered parenterally or intraperitoneally. By
10 way of illustration, solutions of the active compounds as free base or pharmacologically acceptable salts can be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions can also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations generally contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include, for example, sterile
15 aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. Generally, these preparations are sterile and fluid to the extent that easy injectability exists. Preparations should be stable under the conditions of manufacture and storage and should be preserved against the contaminating action of
20 microorganisms, such as bacteria and fungi. Appropriate solvents or dispersion media may contain, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and vegetable oils. The proper fluidity can be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of
25 surfactants. The prevention of the action of microorganisms can be brought about by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example,
30 aluminum monostearate and gelatin.

Sterile injectable solutions may be prepared by incorporating the active compounds in an appropriate amount into a solvent along with any other ingredients (for example as enumerated above) as desired, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which

contains the basic dispersion medium and the desired other ingredients, *e.g.*, as enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation include vacuum-drying and freeze-drying techniques which yield a powder of the active ingredient(s) plus any additional desired ingredient from a previously sterile-filtered solution thereof.

In some embodiments, the compositions of the present disclosure are formulated in a neutral or salt form. Pharmaceutically-acceptable salts include, for example, acid addition salts (formed with the free amino groups of the protein) derived from inorganic acids (*e.g.*, hydrochloric or phosphoric acids, or from organic acids (*e.g.*, acetic, oxalic, tartaric, mandelic, and the like. Salts formed with the free carboxyl groups of the protein can also be derived from inorganic bases (*e.g.*, sodium, potassium, ammonium, calcium, or ferric hydroxides) or from organic bases (*e.g.*, isopropylamine, trimethylamine, histidine, procaine and the like.

Upon formulation, solutions are preferably administered in a manner compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations may easily be administered in a variety of dosage forms such as injectable solutions, drug release capsules and the like. For parenteral administration in an aqueous solution, for example, the solution generally is suitably buffered and the liquid diluent first rendered isotonic for example with sufficient saline or glucose. Such aqueous solutions may be used, for example, for intravenous, intramuscular, subcutaneous and intraperitoneal administration. Preferably, sterile aqueous media are employed as is known to those of skill in the art, particularly in light of the present disclosure. By way of illustration, a single dose may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility, pyrogenicity, general safety and purity standards as required by FDA Office of Biologics standards.

In some embodiments, the Cpfl and gRNAs described herein may be delivered to the patient using adoptive cell transfer (ACT). In adoptive cell transfer, one or more expression constructs are provided *ex vivo* to cells which have originated from the patient (autologous) or from one or more individual(s) other than the patient (allogeneic). The cells are

subsequently introduced or reintroduced into the patient. Thus, in some embodiments, one or more nucleic acids encoding Cpf1 and a guide RNA that targets a dystrophin splice site are provided to a cell ex vivo before the cell is introduced or reintroduced to a patient.

V. Sequence Tables

- 5 The following tables provide exemplary primer sequences and gRNA sequences for use in connection with the compositions and methods disclosed herein.

TABLE C – PRIMER SEQUENCES

	Primer Name		Primer Sequence
Cloning primers for pCpfl-2A-GFP	AgeI-nLbCpfl-F1	F	tttttttcaggttGGaccggtgccaccATGAGCAAGCTGGA (SEQ ID NO: 8)
	nLbCpfl-R1	R	TGGGGTTATAGTAGGCCATCCACTTC (SEQ ID NO: 9)
	nLbCpfl-F2	F	GATGGCCTACTATAACCCAGCG (SEQ ID NO: 10)
	nLbCpfl-R2	R	GGCATAGTCGGGGACATCATATG (SEQ ID NO: 11)
	AgeI-nAsCpfl-F1	F	tttttttcaggttGGaccggtgccaccATGACACAGTTCGAG (SEQ ID NO: 12)
	nAsCpfl-R1	R	TCCTTCTCAGGATTGTTTCAGGTCGTA (SEQ ID NO: 13)
	nAsCpfl-F2	F	CTGAACAATCCTGAGAAGGAGCC (SEQ ID NO: 14)
	nAsCpfl-R2	R	GGCATAGTCGGGGACATCATATG (SEQ ID NO: 11)
	nCpfl-2A-GFP-F	F	ATGATGTCCCCGACTATGCCgaattcGGCAGTGGAGAGGG (SEQ ID NO: 15)
	nCpfl-2A-GFP-R	R	AGCGAGCTCTAGttagaattcCTTGACAG (SEQ ID NO: 16)
In vitro transcription of LbCpfl mRNA	T7-Scaffold-F	F	CACCAGCGCTGCTTAATACGACTCACTATAGGGAAAT (SEQ ID NO: 17)
	T7-Scaffold-R	R	AGTAGCGCTTCTAGACCCTCACTTCCTACTCAG (SEQ ID NO: 18)
	T7-nLb-F1	F	AGAAGAAATATAAGACTCGAGgccaccATGAGCAAGCTGGAGAAGTTTAC (SEQ ID NO: 19)
	T7-nLb-R1	R	TGGGGTTATAGTAGGCCATCC (SEQ ID NO: 20)
	T7-nLb-NLS-F2	F	GATGGCCTACTATAACCCAGCG (SEQ ID NO: 10)
	T7-nLb-NLS-R2	R	CCCCGAGAAGGCAGCGTCGACTTAGGCATAGTCGGGACATCATATG (SEQ ID NO: 21)
	T7-nAs-F1	F	AGAAGAAATATAAGACTCGAGgccaccATGACACAGTTCGAGGGCTTTAC (SEQ ID NO: 22)
	T7-nAs-R1	R	TCCTTCTCAGGATTGTTTCAGGTCGTA (SEQ ID NO: 13)
	T7-nAs-NLS-F2	F	CTGAACAATCCTGAGAAGGAGCC (SEQ ID NO: 14)
	T7-nAs-NLS-R2	R	CCCCGAGAAGGCAGCGTCGACTTAGGCATAGTCGGGACATCATATG (SEQ ID NO: 21)
Human DMD Exon 51 gRNA	nLb-DMD-E51-g1-Top	F	CACCGTAATTTCTACTAAGTGATAGTgCTCCTACTCAGACTGTTACTCTGTTTTTTT (SEQ ID NO: 23)
	nLb-DMD-E51-g1-Bot	R	AAACAAAAAACAGAGTAACAGTCTGAGTAGGAGcATCTACACTTAGTAGAAATTAC (SEQ ID NO: 24)
	nLb-DMD-E51-g2-Top	F	CACCGTAATTTCTACTAAGTGATAGTtaccatgtattgctaacaagtaTTTTTTT (SEQ ID NO: 25)
	nLb-DMD-E51-g2-Bot	R	AAACAAAAAAtactttgttagcaatacatggtaATCTACACTTAGTAGAAATTAC (SEQ ID NO: 26)
	nLb-DMD-E51-g3-Top	F	CACCGTAATTTCTACTAAGTGATAGTattgaagagtaacaatttgagccaTTTTTTT (SEQ ID NO: 27)
	nLb-DMD-E51-g3-Bot	R	AAACAAAAAAtggctcaaatgttactcttcaatATCTACACTTAGTAGAAATTAC (SEQ ID NO: 28)
	nAs-DMD-E51-g1-Top	F	CACCGTAATTTCTACTCTTGATAGTgCTCCTACTCAGACTGTTACTCTGTTTTTTT (SEQ ID NO: 29)
	nAs-DMD-E51-g1-Bot	R	AAACAAAAAACAGAGTAACAGTCTGAGTAGGAGcATCTACAAGAGTAGAAATTAC (SEQ ID NO: 30)
Human DMD Exon 51 T7E1	DMD-E51-T7E1-F1	F	Ttccctggcaaggtctga (SEQ ID NO: 31)
	DMD-E51-T7E1-R1	R	ATCCTCAAGGTCACCCACC (SEQ ID NO: 32)
Human cardiomyocytes RT-PCR	Riken51-RT-PCR-F1	F	CCCAGAAGAGCAAGATAAACTTGA (SEQ ID NO: 1)
	Riken51-RT-PCR-R1	R	CTCTGTTCCAAATCCTGCAATTGT (SEQ ID NO: 33)
Human cardiomyocytes mtDNA copy number qPCR	hmtND1-qF1	F	CGCCACATCTACCATCACCTC (SEQ ID NO: 3)
	hmtND1-qR1	R	CGGCTAGGCTAGAGGTGGCTA (SEQ ID NO: 4)
	hLPL-qF1	F	GAGTATGCAGAGCCCCGAGTC (SEQ ID NO: 5)
	hLPL-qR1	R	TCAACATGCCCAACTGCTTTCTCG (SEQ ID NO: 6)

	Primer Name		Primer Sequence
Mouse Dmd Exon 23 gRNA	nLb-dmd-E23-g1-Top	F	CACCGTAATTTCTACTAAGTGATAGTtagctctgcaaagttctTTGAAAGTTTTTTT (SEQ ID NO: 34)
	nLb-dmd-E23-g1-Bot	R	AAACAAAAAATTTCAAagaactttgcagagcctATCTACACTTAGTAGAAATTAC (SEQ ID NO: 35)

	nLb-dmd-E23-g2-Top	F	CACCGTAATTTCTACTAAGTGTAGATAAAGAGCAACAAATGGCttcaacTTTTTTT (SEQ ID NO: 36)
	nLb-dmd-E23-g2-Bot	R	AAACAAAAAAAgttgaaGCCATTTTGTGCTCTTTATCTACACTTAGTAGAAATTAC (SEQ ID NO: 37)
	nLb-dmd-E23-g2-Top	F	CACCGTAATTTCTACTAAGTGTAGATAAAGAGCAATAAATGGCttcaacTTTTTTT (SEQ ID NO: 38)
	nLb-dmd-E23-g2-Bot	R	AAACAAAAAAAgttgaaGCCATTTTATTGCTCTTTATCTACACTTAGTAGAAATTAC (SEQ ID NO: 39)
	nLb-dmd-E23-g3-Top	F	CACCGTAATTTCTACTAAGTGTAGATAAAGAACTTGCAGAGCctcaaaaTTTTTTT (SEQ ID NO: 40)
	nLb-dmd-E23-g3-Bot	R	AAACAAAAAAAttttgagGCTCTGCAAAGTTCTTTATCTACACTTAGTAGAAATTAC (SEQ ID NO: 41)
	nLb-dmd-I22-g1-Top	F	CACCGTAATTTCTACTAAGTGTAGATctgaatatctatgcattaataactTTTTTTT (SEQ ID NO: 42)
	nLb-dmd-I22-g1-Bot	R	AAACAAAAAAAgtttataatgcatagatattcagATCTACACTTAGTAGAAATTAC (SEQ ID NO: 43)
	nLb-dmd-I22-g2-Top	F	CACCGTAATTTCTACTAAGTGTAGATtattatattacagggcattataTTTTTTT (SEQ ID NO: 44)
	nLb-dmd-I22-g2-Bot	R	AAACAAAAAAAtataatattgcctgtaatataataATCTACACTTAGTAGAAATTAC (SEQ ID NO: 45)
	nLb-dmd-I23-g3-Top	F	CACCGTAATTTCTACTAAGTGTAGATAGGtaagccgaggtttggcctttaTTTTTTT (SEQ ID NO: 46)
	nLb-dmd-I23-g3-Bot	R	AAACAAAAAAAtaaggccaaacctcggttacCTATCTACACTTAGTAGAAATTAC (SEQ ID NO: 47)
	nLb-dmd-I23-g4-Top	F	CACCGTAATTTCTACTAAGTGTAGATcccagagtccttcaaagatattgaTTTTTTT (SEQ ID NO: 48)
	nLb-dmd-I23-g4-Bot	R	AAACAAAAAAAtcaatatctttgaaggactctgggATCTACACTTAGTAGAAATTAC (SEQ ID NO: 49)
In vitro transcription of LbCpf1 gRNA	T7-Lb-dmd-E23-uF	F	GAATTGTAATACGACTCACTATAGGGTAATTTCTACTAAGTGTAGAT (SEQ ID NO: 50)
	T7-Lb-dmd-E23-g1-R	R	CTTTCAAagaactttgcagagcctATCTACACTTAGTAGAAATTA (SEQ ID NO: 51)
	T7-Lb-dmd-E23-mg2-R	R	GttgaaGCCATTTTATTGCTCTTTATCTACACTTAGTAGAAATTA (SEQ ID NO: 52)
	T7-Lb-dmd-E23-g3-R	R	ttttgagGCTCTGCAAAGTTCTTTATCTACACTTAGTAGAAATTA (SEQ ID NO: 53)
	T7-Lb-dmd-I22-g2-R	R	tataatatgccctgtaatataataATCTACACTTAGTAGAAATTACCCTATAGTGAG (SEQ ID NO: 54)
	T7-Lb-dmd-I22-g4-R	R	tcaatatctttgaaggactctgggATCTACACTTAGTAGAAATTACCCTATAGTGAG (SEQ ID NO: 55)
Mouse <i>Dmd</i> Exon 23 T7E1	Dmd-E23-T7E1-F729	F	Gagaaacttctgtgatgtgaggacata (SEQ ID NO: 56)
	Dmd-E23-T7E1-R1	R	CAACCTCGGCTTACCTGAAAT (SEQ ID NO: 57)
	Dmd-E23-T7E1-R729	R	Caatatctttgaaggactctgggtaaa (SEQ ID NO: 58)
	Dmd-E23-T7E1-R3	R	Aattaatagaagtcaatgtagggaagg (SEQ ID NO: 59)

TABLE D – Genomic Target Sequences

Targeted gRNA Exon	Guide #	Strand	Genomic Target Sequence*	PAM	SEQ ID NO.
Human-Exon 51	4	1	tccttttctctttttcctttt	tttt	60
Human-Exon 51	5	1	ctttttctctttttccttttG	tttt	61
Human-Exon 51	6	1	tttttctctttttccttttGC	tttc	62
Human-Exon 51	7	1	tccttttttctcttttGCAAAA	tttt	63
Human-Exon 51	8	1	ctcttttttctcttttGCAAAAA	tttt	64
Human-Exon 51	9	1	ttcttttttctcttttGCAAAAAC	tttc	65
Human-Exon 51	10	1	ttcctttttGCAAAAACCCAAAAT	tttt	66
Human-Exon 51	11	1	tcctttttGCAAAAACCCAAAATA	tttt	67
Human-Exon 51	12	1	ccctttttGCAAAAACCCAAAATAT	tttt	68
Human-Exon 51	13	1	ctttttGCAAAAACCCAAAATATT	tttc	69
Human-Exon 51	14	1	tGCAAAAACCCAAAATATTTTAGC	tttt	70
Human-Exon 51	15	1	GCAAAAACCCAAAATATTTTAGCT	tttt	71
Human-Exon 51	16	1	CAAAAACCCAAAATATTTTAGCTC	tttG	72
Human-Exon 51	17	1	AGCTCCTACTCAGACTGTTACTCT	TTTT	73
Human-Exon 51	18	1	GCTCCTACTCAGACTGTTACTCTG	TTTA	74
Human-Exon 51	19	-1	CTTAGTAACCACAGGTTGTGTCAC	TTTC	75
Human-Exon 51	20	-1	GAGATGGCAGTTTCCTTAGTAACC	TTTG	76
Human-Exon 51	21	-1	TAGTTTGGAGATGGCAGTTTCCTT	TTTC	77
Human-Exon 51	22	-1	TTCTCATACCTTCTGCTTGATGAT	TTTT	78
Human-Exon 51	23	-1	TCATTTTTTCTCATACCTTCTGCT	TTTA	79
Human-Exon 51	24	-1	ATCATTTTTTCTCATACCTTCTGC	TTTT	80
Human-Exon 51	25	-1	AAGAAAAACTTCTGCCAACTTTTA	TTTA	81
Human-Exon 51	26	-1	AAAGAAAAACTTCTGCCAACTTTT	TTTT	82
Human-Exon 51	27	1	TCTTTAAAATGAAGATTTTCCACC	TTTT	83
Human-Exon 51	28	1	CTTTAAAATGAAGATTTTCCACCA	TTTT	84
Human-Exon 51	29	1	TTTAAAATGAAGATTTTCCACCAA	TTTC	85
Human-Exon 51	30	1	AAATGAAGATTTTCCACCAATCAC	TTTA	86
Human-Exon 51	31	1	CCACCAATCACTTTACTCTCCTAG	TTTT	87
Human-Exon 51	32	1	CACCAATCACTTTACTCTCCTAGA	TTTC	88
Human-Exon 51	33	1	CTCTCCTAGACCATTTCACCAG	TTTA	89
Human-Exon 45	1	-1	agaaaagattaaacagtgtgtac	tttg	90
Human-Exon 45	2	-1	tttgagaaaagattaaacagtgtg	TTTa	91
Human-Exon 45	3	-1	atttgagaaaagattaaacagtgt	TTTT	92
Human-Exon 45	4	-1	Tatttgagaaaagattaaacagtg	TTTT	93
Human-Exon 45	5	1	atcttttctcaatAAAAAGACAT	ttta	94
Human-Exon 45	6	1	ctcaaatAAAAAGACATGGGGCTT	tttt	95
Human-Exon 45	7	1	tcaaatAAAAAGACATGGGGCTTC	tttc	96

Human-Exon 45	8	1	TGTTTTGCCTTTTTGGTATCTTAC	TTTT	97
Human-Exon 45	9	1	GTTTTGCCTTTTTGGTATCTTACA	TTTT	98
Human-Exon 45	10	1	TTTTGCCTTTTTGGTATCTTACAG	TTTG	99
Human-Exon 45	11	1	GCCTTTTTGGTATCTTACAGGAAC	TTTT	100
Human-Exon 45	12	1	CCTTTTTGGTATCTTACAGGAACT	TTTG	101
Human-Exon 45	13	1	TGGTATCTTACAGGAACTCCAGGA	TTTT	102
Human-Exon 45	14	1	GGTATCTTACAGGAACTCCAGGAT	TTTT	103
Human-Exon 45	15	-1	AGGATTGCTGAATTATTTCTTCCC	TTTG	104
Human-Exon 45	16	-1	GAGGATTGCTGAATTATTTCTTCC	TTTT	105
Human-Exon 45	17	-1	TGAGGATTGCTGAATTATTTCTTC	TTTT	106
Human-Exon 45	18	-1	CTGTAGAATACTGGCATCTGTTTT	TTTC	107
Human-Exon 45	19	-1	CCTGTAGAATACTGGCATCTGTTT	TTTT	108
Human-Exon 45	20	-1	TCCTGTAGAATACTGGCATCTGTT	TTTT	109
Human-Exon 45	21	-1	CAGACCTCCTGCCACCGCAGATTC	TTTG	110
Human-Exon 45	22	-1	TGTCTGACAGCTGTTTGCAGACCT	TTTC	111
Human-Exon 45	23	-1	CTGTCTGACAGCTGTTTGCAGACC	TTTT	112
Human-Exon 45	24	-1	TCTGTCTGACAGCTGTTTGCAGAC	TTTT	113
Human-Exon 45	25	-1	TTCTGTCTGACAGCTGTTTGCAGA	TTTT	114
Human-Exon 45	26	-1	ATTCCTATTAGATCTGTCGCCCTA	TTTC	115
Human-Exon 45	27	-1	CATTCCTATTAGATCTGTCGCCCT	TTTT	116
Human-Exon 45	28	1	AGCAGACTTTTTAAGCTTTCTTTA	TTTT	117
Human-Exon 45	29	1	GCAGACTTTTTAAGCTTTCTTTAG	TTTA	118
Human-Exon 45	30	1	TAAGCTTTCTTTAGAAGAATATTT	TTTT	119
Human-Exon 45	31	1	AAGCTTTCTTTAGAAGAATATTTT	TTTT	120
Human-Exon 45	32	1	AGCTTTCTTTAGAAGAATATTTCA	TTTA	121
Human-Exon 45	33	1	TTAGAAGAATATTTTCATGAGAGA	TTTC	122
Human-Exon 45	34	1	GAAGAATATTTTCATGAGAGATTAT	TTTA	123
Human-Exon 44	1	1	TCAGTATAACCAAAAAATATACGC	TTTG	124
Human-Exon 44	2	1	acataatccatctatTTTTcttga	tttt	125
Human-Exon 44	3	1	cataatccatctatTTTTcttgat	ttta	126
Human-Exon 44	4	1	tcttgatccatgcttttACCTG	tttt	127
Human-Exon 44	5	1	cttgatccatgcttttACCTGC	tttt	128
Human-Exon 44	6	1	ttgatccatgcttttACCTGCA	tttc	129
Human-Exon 44	7	-1	TCAACAGATCTGTCAAATCGCCTG	TTTC	130
Human-Exon 44	8	1	ACCTGCAGGCGATTTGACAGATCT	tttt	131
Human-Exon 44	9	1	CCTGCAGGCGATTTGACAGATCTG	tttA	132
Human-Exon 44	10	1	ACAGATCTGTTGAGAAATGGCGGC	TTTG	133
Human-Exon 44	11	-1	TATCATAATGAAAACGCCGCCATT	TTTA	134
Human-Exon 44	12	1	CATTATGATATAAAGATATTTAAT	TTTT	135
Human-Exon 44	13	-1	TATTTAGCATGTTCCCAATTCTCA	TTTG	136
Human-Exon 44	14	-1	GAAAAAACAAATCAAAGACTTACC	TTTC	137
Human-Exon 44	15	1	ATTTGTTTTTTTCGAAATTGTATTT	TTTG	138

Human-Exon 44	16	1	TTTTTTCGAAATTGTATTTATCTT	TTTG	139
Human-Exon 44	17	1	TTCGAAATTGTATTTATCTTCAGC	TTTT	140
Human-Exon 44	18	1	TCGAAATTGTATTTATCTTCAGCA	TTTT	141
Human-Exon 44	19	1	CGAAATTGTATTTATCTTCAGCAC	TTTT	142
Human-Exon 44	20	1	GAAATTGTATTTATCTTCAGCACA	TTTC	143
Human-Exon 44	21	-1	AGAAGTTAAAGAGTCCAGATGTGC	TTTA	144
Human-Exon 44	22	1	TCTTCAGCACATCTGGACTCTTTA	TTTA	145
Human-Exon 44	23	-1	CATCACCTTCAGAACCTGATCTT	TTTC	146
Human-Exon 44	24	1	ACTTCTTAAAGATCAGGTTCTGAA	TTTA	147
Human-Exon 44	25	1	GACTGTTGTTGTCATCATTATATT	TTTT	148
Human-Exon 44	26	1	ACTGTTGTTGTCATCATTATATTA	TTTG	149
Human-Exon 53	1	-1	AACTAGAATAAAAGGAAAAATAAA	TTTC	150
Human-Exon 53	2	1	CTACTATATATTTATTTTCCTTT	TTTA	151
Human-Exon 53	3	1	TTTTTCCTTTTATTCTAGTTGAAA	TTTA	152
Human-Exon 53	4	1	TCCTTTTATTCTAGTTGAAAGAAT	TTTT	153
Human-Exon 53	5	1	CCTTTTATTCTAGTTGAAAGAATT	TTTT	154
Human-Exon 53	6	1	CTTTTATTCTAGTTGAAAGAATTC	TTTC	155
Human-Exon 53	7	1	ATTCTAGTTGAAAGAATTCAGAAT	TTTT	156
Human-Exon 53	8	1	TTCTAGTTGAAAGAATTCAGAATC	TTTA	157
Human-Exon 53	9	-1	ATTCAACTGTTGCCTCCGTTCTG	TTTC	158
Human-Exon 53	10	-1	ACATTTCAATCAACTGTTGCCTCC	TTTA	159
Human-Exon 53	11	-1	CTTTTGGATTGCATCTACTGTATA	TTTT	160
Human-Exon 53	12	-1	TGTGATTTTCTTTTGGATTGCATC	TTTC	161
Human-Exon 53	13	-1	ATACTAACCTTGGTTTCTGTGATT	TTTG	162
Human-Exon 53	14	-1	AAAAGGTATCTTTGATACTAACCT	TTTA	163
Human-Exon 53	15	-1	AAAAGGTATCTTTGATACTAACC	TTTT	164
Human-Exon 53	16	-1	TTTTAAAAGGTATCTTTGATACT	TTTA	165
Human-Exon 53	17	-1	ATTTTAAAAGGTATCTTTGATAC	TTTT	166
Human-Exon 46	1	-1	TTAATGCAAACCTGGGACACAAACA	TTTG	167
Human-Exon 46	2	1	TAAATTGCCATGTTTGTGTCCCAG	TTTT	168
Human-Exon 46	3	1	AAATTGCCATGTTTGTGTCCCAGT	TTTT	169
Human-Exon 46	4	1	AATTGCCATGTTTGTGTCCCAGTT	TTTA	170
Human-Exon 46	5	1	TGTCCCAGTTTGCATTAACAAATA	TTTG	171
Human-Exon 46	6	-1	CAACATAGTTCTCAAACATTTGT	tttC	172
Human-Exon 46	7	-1	CCAACATAGTTCTCAAACATTTG	tttt	173
Human-Exon 46	8	-1	tCCAACATAGTTCTCAAACATTT	tttt	174
Human-Exon 46	9	-1	tttCCAACATAGTTCTCAAACAT	tttt	175
Human-Exon 46	10	-1	ttttCCAACATAGTTCTCAAACAT	tttt	176
Human-Exon 46	11	-1	tttttCCAACATAGTTCTCAAACAT	tttt	177
Human-Exon 46	12	1	CATTAACAAATAGTTTGAGAACTA	TTTG	178
Human-Exon 46	13	1	AGAACTATGTTGGaaaaaaaTA	TTTG	179
Human-Exon 46	14	-1	GTTCTTCTAGCCTGGAGAAAGAAG	TTTT	180

Human-Exon 46	15	1	ATTCTTCTTTCTCCAGGCTAGAAG	TTTT	181
Human-Exon 46	16	1	TTCTTCTTTCTCCAGGCTAGAAGA	TTTA	182
Human-Exon 46	17	1	TCCAGGCTAGAAGAACAAAAGAAT	TTTC	183
Human-Exon 46	18	-1	AAATTCTGACAAGATATTCTTTTG	TTTG	184
Human-Exon 46	19	-1	CTTTTAGTTGCTGCTCTTTTCCAG	TTTT	185
Human-Exon 46	20	-1	AGAAAATAAAATTACCTTGACTTG	TTTG	186
Human-Exon 46	21	-1	TGCAAGCAGGCCCTGGGGGATTG	TTTA	187
Human-Exon 46	22	1	ATTTTCTCAAATCCCCCAGGGCCT	TTTT	188
Human-Exon 46	23	1	TTTTCTCAAATCCCCCAGGGCCTG	TTTA	189
Human-Exon 46	24	1	CTCAAATCCCCCAGGGCCTGCTTG	TTTT	190
Human-Exon 46	25	1	TCAAATCCCCCAGGGCCTGCTTGC	TTTC	191
Human-Exon 46	26	1	TTAATTCAATCATTGGTTTTCTGC	TTTT	192
Human-Exon 46	27	1	TAATTCAATCATTGGTTTTCTGCC	TTTT	193
Human-Exon 46	28	1	AATTCAATCATTGGTTTTCTGCCC	TTTT	194
Human-Exon 46	29	1	ATTCAATCATTGGTTTTCTGCCCA	TTTA	195
Human-Exon 46	30	-1	GCAAGGA ACTATGAATAACCTAAT	TTTA	196
Human-Exon 46	31	1	CTGCCCATTAGGTTATTCATAGTT	TTTT	197
Human-Exon 46	32	1	TGCCCATTAGGTTATTCATAGTTC	TTTC	198
Human-Exon 52	1	-1	TAGAAAACAATTTAACAGGAAATA	TTTA	199
Human-Exon 52	2	1	CTGTTAAATTGTTTTCTATAAACC	TTTC	200
Human-Exon 52	3	-1	GAAATAAAAAAAGATGTTACTGTAT	TTTA	201
Human-Exon 52	4	-1	AGAAATAAAAAAAGATGTTACTGTA	TTTT	202
Human-Exon 52	5	1	CTATAAACCCCTTATACAGTAACAT	TTTT	203
Human-Exon 52	6	1	TATAAACCCCTTATACAGTAACATC	TTTC	204
Human-Exon 52	7	1	TTATTTCTAAAAGTGTTTTGGCTG	TTTT	205
Human-Exon 52	8	1	TATTTCTAAAAGTGTTTTGGCTGG	TTTT	206
Human-Exon 52	9	1	ATTTCTAAAAGTGTTTTGGCTGGT	TTTT	207
Human-Exon 52	10	1	TTTCTAAAAGTGTTTTGGCTGGTC	TTTA	208
Human-Exon 52	11	1	TAAAAGTGTTTTGGCTGGTCTCAC	TTTC	209
Human-Exon 52	12	-1	CATAATACAAAGTAAAGTACAATT	TTTA	210
Human-Exon 52	13	-1	ACATAATACAAAGTAAAGTACAAT	TTTT	211
Human-Exon 52	14	1	GGCTGGTCTCACAATTGTACTTTA	TTTT	212
Human-Exon 52	15	1	GCTGGTCTCACAATTGTACTTTAC	TTTG	213
Human-Exon 52	16	1	CTTTGTATTATGTAAAAGGAATAC	TTTA	214
Human-Exon 52	17	1	TATTATGTAAAAGGAATACACAAC	TTTG	215
Human-Exon 52	18	1	TTCTTACAGGCAACAATGCAGGAT	TTTG	216
Human-Exon 52	19	1	GAACAGAGGCGTCCCCAGTTGGAA	TTTG	217
Human-Exon 52	20	-1	GGCAGCGGTAATGAGTTCTTCCAA	TTTG	218
Human-Exon 52	21	-1	TCAAATTTTGGGCAGCGGTAATGA	TTTT	219
Human-Exon 52	22	1	AAAAACAAGACCAGCAATCAAGAG	TTTG	220
Human-Exon 52	23	-1	TGTGTCCCATGCTTGTTAAAAAAC	TTTG	221
Human-Exon 52	24	1	TTACAAGCATGGGACACACAAAG	TTTT	222

Human-Exon 52	25	1	TAACAAGCATGGGACACACAAAGC	TTTT	223
Human-Exon 52	26	1	AACAAGCATGGGACACACAAAGCA	TTTT	224
Human-Exon 52	27	1	ACAAGCATGGGACACACAAAGCAA	TTTA	225
Human-Exon 52	28	-1	TTGAAACTTGTTCATGCATCTTGCT	TTTA	226
Human-Exon 52	29	-1	ATTGAAACTTGTTCATGCATCTTGC	TTTT	227
Human-Exon 52	30	-1	TATTGAAACTTGTTCATGCATCTTG	TTTT	228
Human-Exon 52	31	1	AATAAAAACCTTAAGTTCATATATC	TTTC	229
Human-Exon 50	1	-1	GTGAATATATTATTGGATTTCTAT	TTTG	230
Human-Exon 50	2	-1	AAGATAATTCATGAACATCTTAAT	TTTG	231
Human-Exon 50	3	-1	ACAGAAAAGCATACACATTACTTA	TTTA	232
Human-Exon 50	4	1	CTGTAAAGAGGAAGTTAGAAGAT	TTTT	233
Human-Exon 50	5	1	TGTTAAAGAGGAAGTTAGAAGATC	TTTC	234
Human-Exon 50	6	-1	CCGCCTTCCACTCAGAGCTCAGAT	TTTA	235
Human-Exon 50	7	-1	CCCTCAGCTCTTGAAGTAAACGGT	TTTG	236
Human-Exon 50	8	1	CTTCAAGAGCTGAGGGCAAAGCAG	TTTA	237
Human-Exon 50	9	-1	AACAAATAGCTAGAGCCAAAGAGA	TTTG	238
Human-Exon 50	10	-1	GAACAAATAGCTAGAGCCAAAGAG	TTTT	239
Human-Exon 50	11	1	GCTCTAGCTATTTGTTCAAAGTG	TTTG	240
Human-Exon 50	12	1	TTCAAAGTGCAACTATGAAGTGA	TTTG	241
Human-Exon 50	13	-1	TCTCTACCCAGTCATCACTTCAT	TTTC	242
Human-Exon 50	14	-1	CTCTCTACCCAGTCATCACTTCA	TTTT	243
Human-Exon 43	1	1	tatatatatatatTTTTCTCTT	TTTG	244
Human-Exon 43	2	1	TCTCTTTCTATAGACAGCTAATTC	tTTT	245
Human-Exon 43	3	1	CTCTTTCTATAGACAGCTAATTCA	TTTT	246
Human-Exon 43	4	-1	AAACAGTAAAAAAATGAATTAGCT	TTTA	247
Human-Exon 43	5	1	TCTTTCTATAGACAGCTAATTCAT	TTTC	248
Human-Exon 43	6	-1	AAAACAGTAAAAAAATGAATTAGC	TTTT	249
Human-Exon 43	7	1	TATAGACAGCTAATTCATTTTTTT	TTTC	250
Human-Exon 43	8	-1	TATTCTGTAATATAAAAATTTTAA	TTTA	251
Human-Exon 43	9	-1	ATATTCTGTAATATAAAAATTTTA	TTTT	252
Human-Exon 43	10	1	TTTACTGTTTTAAAATTTTTATAT	TTTT	253
Human-Exon 43	11	1	TTACTGTTTTAAAATTTTTATATT	TTTT	254
Human-Exon 43	12	1	TACTGTTTTAAAATTTTTATATTA	TTTT	255
Human-Exon 43	13	1	ACTGTTTTAAAATTTTTATATTAC	TTTT	256
Human-Exon 43	14	1	CTGTTTTAAAATTTTTATATTACA	TTTA	257
Human-Exon 43	15	1	AAAATTTTTATATTACAGAATATA	TTTT	258
Human-Exon 43	16	1	AAATTTTTATATTACAGAATATAA	TTTA	259
Human-Exon 43	17	-1	TTGTAGACTATCTTTTATATTCTG	TTTG	260
Human-Exon 43	18	1	TATATTACAGAATATAAAAAGATAG	TTTT	261
Human-Exon 43	19	1	ATATTACAGAATATAAAAAGATAGT	TTTT	262
Human-Exon 43	20	1	TATTACAGAATATAAAAAGATAGTC	TTTA	263
Human-Exon 43	21	-1	CAATGCTGCTGTCTTCTTGCTATG	TTTG	264

Human-Exon 43	22	1	CAATGGGAAAAAGTTAACAAAATG	TTTC	265
Human-Exon 43	23	-1	TGCAAGTATCAAGAAAAATATATG	TTTC	266
Human-Exon 43	24	1	TCTTGATACTTGCAGAAATGATTT	TTTT	267
Human-Exon 43	25	1	CTTGATACTTGCAGAAATGATTTG	TTTT	268
Human-Exon 43	26	1	TTGATACTTGCAGAAATGATTTGT	TTTC	269
Human-Exon 43	27	1	TTTTCAGGGAAGTGTAGAATTTAT	TTTG	270
Human-Exon 43	28	-1	CATGGAGGGTACTGAAATAAATTC	TTTC	271
Human-Exon 43	29	-1	CCATGGAGGGTACTGAAATAAATT	TTTT	272
Human-Exon 43	30	1	CAGGGAAGTGTAGAATTTATTTCA	TTTT	273
Human-Exon 43	31	-1	TCCATGGAGGGTACTGAAATAAAT	TTTT	274
Human-Exon 43	32	1	AGGGAAGTGTAGAATTTATTTTCAG	TTTC	275
Human-Exon 43	33	-1	TTCCATGGAGGGTACTGAAATAAA	TTTT	276
Human-Exon 43	34	-1	CCTGTCTTTTTTCCATGGAGGGTA	TTTC	277
Human-Exon 43	35	-1	CCCTGTCTTTTTTCCATGGAGGGT	TTTT	278
Human-Exon 43	36	-1	TCCCTGTCTTTTTTCCATGGAGGG	TTTT	279
Human-Exon 43	37	1	TTTCAGTACCCTCCATGGAAAAAA	TTTA	280
Human-Exon 43	38	1	AGTACCCTCCATGGAAAAAAGACA	TTTC	281
Human-Exon 6	1	1	AGTTTGCATGGTTCTTGCTCAAGG	TTTA	282
Human-Exon 6	2	-1	ATAAGAAAATGCATTCCTTGAGCA	TTTC	283
Human-Exon 6	3	-1	CATAAGAAAATGCATTCCTTGAGC	TTTT	284
Human-Exon 6	4	1	CATGGTTCTTGCTCAAGGAATGCA	TTTG	285
Human-Exon 6	5	-1	ACCTACATGTGGAAATAAATTTTC	TTTG	286
Human-Exon 6	6	-1	GACCTACATGTGGAAATAAATTTT	TTTT	287
Human-Exon 6	7	-1	TGACCTACATGTGGAAATAAATTT	TTTT	288
Human-Exon 6	8	1	CTTATGAAAATTTATTTCCACATG	TTTT	289
Human-Exon 6	9	1	TTATGAAAATTTATTTCCACATGT	TTTC	290
Human-Exon 6	10	-1	ATTACATTTTTTGACCTACATGTGG	TTTC	291
Human-Exon 6	11	-1	CATTACATTTTTTGACCTACATGTG	TTTT	292
Human-Exon 6	12	-1	TCATTACATTTTTTGACCTACATGT	TTTT	293
Human-Exon 6	13	1	TTCCACATGTAGGTCAAAAATGT	TTTA	294
Human-Exon 6	14	1	CACATGTAGGTCAAAAATGTAATG	TTTC	295
Human-Exon 6	15	-1	TTGCAATCCAGCCATGATATTTTT	TTTG	296
Human-Exon 6	16	-1	ACTGTTGGTTTGTTGCAATCCAGC	TTTC	297
Human-Exon 6	17	-1	CACTGTTGGTTTGTTGCAATCCAG	TTTT	298
Human-Exon 6	18	1	AATGCTCTCATCCATAGTCATAGG	TTTG	299
Human-Exon 6	19	-1	ATGTCTCAGTAATCTTCTTACCTA	TTTA	300
Human-Exon 6	20	-1	CAAGTTATTTAATGTCTCAGTAAT	TTTA	301
Human-Exon 6	21	-1	ACAAGTTATTTAATGTCTCAGTAA	TTTT	302
Human-Exon 6	22	1	GACTCTGATGACATATTTTTCCCC	TTTA	303
Human-Exon 6	23	1	TCCCCAGTATGGTTCCAGATCATG	TTTT	304
Human-Exon 6	24	1	CCCCAGTATGGTTCCAGATCATGT	TTTT	305
Human-Exon 6	25	1	CCCAGTATGGTTCCAGATCATGTC	TTTC	306

Human-Exon 7	1	1	TATTTGTCTTtgtgtatgtgtga	TTTA	307
Human-Exon 7	2	1	TCTTtgtgtatgtgtatgtga	TTTG	308
Human-Exon 7	3	1	tgatgtgtatgtgtatgtgtt	TTtg	309
Human-Exon 7	4	1	AGGCCAGACCTATTTGACTGGAAT	ttTT	310
Human-Exon 7	5	1	GGCCAGACCTATTTGACTGGAATA	tTTA	311
Human-Exon 7	6	1	ACTGGAATAGTGTGGTTTGCCAGC	TTTG	312
Human-Exon 7	7	1	CCAGCAGTCAGCCACACAACGACT	TTTG	313
Human-Exon 7	8	-1	TCTATGCCTAATTGATATCTGGCG	TTTC	314
Human-Exon 7	9	-1	CCAACCTTCAGGATCGAGTAGTTT	TTTA	315
Human-Exon 7	10	1	TGGACTACCACTGCTTTTAGTATG	TTTC	316
Human-Exon 7	11	1	AGTATGGTAGAGTTTAATGTTTTTC	TTTT	317
Human-Exon 7	12	1	GTATGGTAGAGTTTAATGTTTTCA	TTTA	318
Human-Exon 8	1	-1	AGACTCTAAAAGGATAATGAACAA	TTTG	319
Human-Exon 8	2	1	ACTTTGATTTGTTTCATTATCCTTT	TTTA	320
Human-Exon 8	3	-1	TATATTTGAGACTCTAAAAGGATA	TTTC	321
Human-Exon 8	4	1	ATTTGTTTCATTATCCTTTTAGAGT	TTTG	322
Human-Exon 8	5	-1	GTTTCTATATTTGAGACTCTAAA	TTTG	323
Human-Exon 8	6	-1	GGTTTCTATATTTGAGACTCTAAA	TTTT	324
Human-Exon 8	7	-1	TGGTTTCTATATTTGAGACTCTAA	TTTT	325
Human-Exon 8	8	1	TTCATTATCCTTTTAGAGTCTCAA	TTTG	326
Human-Exon 8	9	1	AGAGTCTCAAATATAGAAACCAAA	TTTT	327
Human-Exon 8	10	1	GAGTCTCAAATATAGAAACCAAAA	TTTA	328
Human-Exon 8	11	-1	CACTTCCTGGATGGCTTCAATGCT	TTTC	329
Human-Exon 8	12	1	GCCTCAACAAGTGAGCATTGAAGC	TTTT	330
Human-Exon 8	13	1	CCTCAACAAGTGAGCATTGAAGCC	TTTG	331
Human-Exon 8	14	-1	GGTGGCCTTGGCAACATTTCCACT	TTTA	332
Human-Exon 8	15	-1	GTCACTTTAGGTGGCCTTGGCAAC	TTTA	333
Human-Exon 8	16	-1	ATGATGTAAGTGAATGTTCTTC	TTTG	334
Human-Exon 8	17	-1	CCTGTTGAGAATAGTGCATTTGAT	TTTA	335
Human-Exon 8	18	1	CAGTTACATCATCAAATGCACTAT	TTTT	336
Human-Exon 8	19	1	AGTTACATCATCAAATGCACTATT	TTTC	337
Human-Exon 8	20	-1	CACACTTTACCTGTTGAGAATAGT	TTTA	338
Human-Exon 8	21	1	CTGTTTTATATGCATTTTTAGGTA	TTTT	339
Human-Exon 8	22	1	TGTTTTATATGCATTTTTAGGTAT	TTTC	340
Human-Exon 8	23	1	ATATGCATTTTTAGGTATTACGTG	TTTT	341
Human-Exon 8	24	1	TATGCATTTTTAGGTATTACGTGC	TTTA	342
Human-Exon 8	25	1	TAGGTATTACGTGCACatatatat	TTTT	343
Human-Exon 8	26	1	AGGTATTACGTGCACatatatata	TTTT	344
Human-Exon 8	27	1	GGTATTACGTGCACatatatatat	TTTA	345
Human-Exon 55	1	-1	AGCAACAACCTATAATATTGTGCAG	TTTA	346
Human-Exon 55	2	1	GTTCTCCATCTTTCTCTTTTAT	TTTA	347
Human-Exon 55	3	1	TCTTTTTATGGAGTTCAGGTG	TTTC	348

Human-Exon 55	4	1	TATGGAGTTCAGGTGCACCAT	TTTT	349
Human-Exon 55	5	1	ATGGAGTTCAGGTGCACCATT	TTTT	350
Human-Exon 55	6	1	TGGAGTTCAGGTGCACCATTC	TTTA	351
Human-Exon 55	7	1	ATAATTGCATCTGAACATTTGGTC	TTTA	352
Human-Exon 55	8	1	GTCCTTTGCAGGGTGAGTGAGCGA	TTTG	353
Human-Exon 55	9	-1	TTCCAAAGCAGCCTCTCGCTCACT	TTTC	354
Human-Exon 55	10	1	CAGGGTGAGTGAGCGAGAGGCTGC	TTTG	355
Human-Exon 55	11	1	GAAGAACTCATAGATTACTGCAA	TTTG	356
Human-Exon 55	12	-1	CAGGTCCAGGGGGAAGTGTTCAG	TTTC	357
Human-Exon 55	13	-1	CCAGGTCCAGGGGGAAGTGTTCAG	TTTT	358
Human-Exon 55	14	-1	AGCTTCTGTAAGCCAGGCAAGAAA	TTTC	359
Human-Exon 55	15	1	TTGCCTGGCTTACAGAAGCTGAAA	TTTC	360
Human-Exon 55	16	-1	CTTACGGGTAGCATCCTGTAGGAC	TTTC	361
Human-Exon 55	17	-1	CTCCCTTGGAGTCTTCTAGGAGCC	TTTA	362
Human-Exon 55	18	-1	ACTCCCTTGGAGTCTTCTAGGAGC	TTTT	363
Human-Exon 55	19	-1	ATCAGCTCTTTTACTCCCTTGGAG	TTTC	364
Human-Exon 55	20	1	CGCTTTAGCACTCTTGTGGATCCA	TTTC	365
Human-Exon 55	21	1	GCACTCTTGTGGATCCAATTGAAC	TTTA	366
Human-Exon 55	22	-1	TCCCTGGCTTGTGAGTTACAAGTA	TTTG	367
Human-Exon 55	23	-1	GTCCCTGGCTTGTGAGTTACAAGT	TTTT	368
Human-Exon 55	24	-1	TTTTGTCCCTGGCTTGTGAGTTAC	TTTG	369
Human-Exon 55	25	-1	GTTTTGTCCCTGGCTTGTGAGTTA	TTTT	370
Human-Exon 55	26	1	TACTTGTAAGTACAAGCCAGGGA	TTTG	371
Human-G1-exon51		1	gCTCCTACTCAGACTGTTACTCTG	TTTA	372
Human-G2-exon51		1	taccatgtattgctaacaaga	TTTC	373
Human-G3-exon51		-1	attgaagagtaacaattgagcca	TTTA	374
mouse-Exon23-G1		1	aggctctgcaaagttctTTGAAAG	TTTG	375
mouse-Exon23-G2		1	AAAGAGCAACAAAATGGCttcaac	TTTG	376
mouse-Exon23-G3		1	AAAGAGCAATAAAAATGGCttcaac	TTTG	377
mouse-Exon23-G4		-1	AAAGAACTTTGCAGAGCctcaaaa	TTTC	378
mouse-Exon23-G5		-1	ctgaatatctatgcattaataact	TTTA	379
mouse-Exon23-G6		-1	tattatattacagggcatattata	TTTC	380

mouse-Exon23-G7		1	Aggtaagccgaggttggccttta	TTTC	381
mouse-Exon23-G8		1	cccagagtccttcaaagatattga	TTTA	382

* In this table, upper case letters represent nucleotides that align to the exon sequence of the gene. Lower case letters represent nucleotides that align to the intron sequence of the gene.

TABLE E – gRNA sequences

5

Targeted gRNA Exon	Guide #	Strand	gRNA sequence *	PAM	SEQ ID NO.
Human-Exon 51	4	1	aaaaaggaaaaagaagaaaaaga	tttt	448
Human-Exon 51	5	1	Caaaaaggaaaaagaagaaaaag	tttt	449
Human-Exon 51	6	1	GCaaaaaggaaaaagaagaaaaa	tttc	450
Human-Exon 51	7	1	UUUUGCaaaaaggaaaaagaaga	tttt	451
Human-Exon 51	8	1	UUUUUGCaaaaaggaaaaagaag	tttt	452
Human-Exon 51	9	1	GUUUUUGCaaaaaggaaaaagaa	tttc	453
Human-Exon 51	10	1	AUUUUGGGUUUUUGCaaaaaggaa	tttt	454
Human-Exon 51	11	1	UAUUUUGGGUUUUUGCaaaaagga	tttt	455
Human-Exon 51	12	1	AUAUUUUGGGUUUUUGCaaaaagg	tttt	456
Human-Exon 51	13	1	AAUAUUUUGGGUUUUUGCaaaaag	tttc	457
Human-Exon 51	14	1	GCUAAAAUAUUUUGGGUUUUUGCa	tttt	458
Human-Exon 51	15	1	AGCUAAAAUAUUUUGGGUUUUUGC	tttt	459
Human-Exon 51	16	1	GAGCUAAAAUAUUUUGGGUUUUUG	tttG	460
Human-Exon 51	17	1	AGAGUAAACAGUCUGAGUAGGAGCU	TTTT	461
Human-Exon 51	18	1	CAGAGUAAACAGUCUGAGUAGGAGC	TTTA	462
Human-Exon 51	19	-1	GUGACACAACCUGUGGUUACUAAG	TTTC	463
Human-Exon 51	20	-1	GGUUACUAAGGAAACUGCCAUCU	TTTG	464
Human-Exon 51	21	-1	AAGGAAACUGCCAUCUCCAAACUA	TTTC	465
Human-Exon 51	22	-1	AUCAUCAAGCAGAAGGUAUGAGAA	TTTT	466
Human-Exon 51	23	-1	AGCAGAAGGUAUGAGAAAAAUGA	TTTA	467
Human-Exon 51	24	-1	GCAGAAGGUAUGAGAAAAAUGAU	TTTT	468
Human-Exon 51	25	-1	UAAAAGUUGGCAGAAGUUUUUCUU	TTTA	469
Human-Exon 51	26	-1	AAAAGUUGGCAGAAGUUUUUCUUU	TTTT	470
Human-Exon 51	27	1	GGUGGAAAAUCUUCAUUUUAAAGA	TTTT	471
Human-Exon 51	28	1	UGGUGGAAAAUCUUCAUUUUAAAG	TTTT	472
Human-Exon 51	29	1	UUGGUGGAAAAUCUUCAUUUUAAA	TTTC	473
Human-Exon 51	30	1	GUGAUUGGUGGAAAAUCUUCAUUU	TTTA	474
Human-Exon 51	31	1	CUAGGAGAGUAAAGUGAUUGGUGG	TTTT	475

Human-Exon 51	32	1	UCUAGGAGAGUAAAGUGAUUGGUG	TTTC	476
Human-Exon 51	33	1	CUGGUGGGAAAUGGUCUAGGAGA	TTTA	477
Human-Exon 45	1	-1	guagcacacuguuuaaucuuuucu	tttg	478
Human-Exon 45	2	-1	cacacuguuuaaucuuuucucaaa	TTTa	479
Human-Exon 45	3	-1	acacuguuuaaucuuuucucaaa	TTTT	480
Human-Exon 45	4	-1	cacuguuuaaucuuuucucaaaA	TTTT	481
Human-Exon 45	5	1	AUGUCUUUUUauuugagaaaagau	ttta	482
Human-Exon 45	6	1	AAGCCCCAUGUCUUUUUauuugag	tttt	483
Human-Exon 45	7	1	GAAGCCCCAUGUCUUUUUauuuga	tttc	484
Human-Exon 45	8	1	GUAAGAUACCAAAAAGGCAAAACA	TTTT	485
Human-Exon 45	9	1	UGUAAGAUACCAAAAAGGCAAAAC	TTTT	486
Human-Exon 45	10	1	CUGUAAGAUACCAAAAAGGCAAAA	TTTG	487
Human-Exon 45	11	1	GUUCCUGUAAGAUACCAAAAAGGC	TTTT	488
Human-Exon 45	12	1	AGUUCCUGUAAGAUACCAAAAAGG	TTTG	489
Human-Exon 45	13	1	UCCUGGAGUUCUGUAAGAUACCA	TTTT	490
Human-Exon 45	14	1	AUCCUGGAGUUCUGUAAGAUACC	TTTT	491
Human-Exon 45	15	-1	GGGAAGAAUAAUUCAGCAAUCCU	TTTG	492
Human-Exon 45	16	-1	GGAAGAAUAAUUCAGCAAUCCUC	TTTT	493
Human-Exon 45	17	-1	GAAGAAUAAUUCAGCAAUCCUCA	TTTT	494
Human-Exon 45	18	-1	AAAACAGAUGCCAGUAUUCUACAG	TTTC	495
Human-Exon 45	19	-1	AAACAGAUGCCAGUAUUCUACAGG	TTTT	496
Human-Exon 45	20	-1	AACAGAUGCCAGUAUUCUACAGGA	TTTT	497
Human-Exon 45	21	-1	GAAUCUGCGGUGGCAGGAGGUCUG	TTTG	498
Human-Exon 45	22	-1	AGGUCUGCAAACAGCUGUCAGACA	TTTC	499
Human-Exon 45	23	-1	GGUCUGCAAACAGCUGUCAGACAG	TTTT	500
Human-Exon 45	24	-1	GUCUGCAAACAGCUGUCAGACAGA	TTTT	501
Human-Exon 45	25	-1	UCUGCAAACAGCUGUCAGACAGAA	TTTT	502
Human-Exon 45	26	-1	UAGGGCGACAGAUCUAAUAGGAAU	TTTC	503
Human-Exon 45	27	-1	AGGGCGACAGAUCUAAUAGGAAUG	TTTT	504
Human-Exon 45	28	1	UAAAGAAAGCUUAAAAAGUCUGCU	TTTT	505
Human-Exon 45	29	1	CUAAAGAAAGCUUAAAAAGUCUGC	TTTA	506
Human-Exon 45	30	1	AAAUUUUCUUCUAAAGAAAGCUUA	TTTT	507
Human-Exon 45	31	1	GAAAUUUUCUUCUAAAGAAAGCUU	TTTT	508
Human-Exon 45	32	1	UGAAAUUUUCUUCUAAAGAAAGCU	TTTA	509
Human-Exon 45	33	1	UCUCUCAUGAAAUUUUCUUCUAAA	TTTC	510
Human-Exon 45	34	1	AUAAUCUCUCAUGAAAUUUUCUUC	TTTA	511
Human-Exon 44	1	1	GCGUAUUUUUUUGGUUAUACUGA	TTTG	512
Human-Exon 44	2	1	ucaagaaaaauagauggauuaugu	tttt	513
Human-Exon 44	3	1	aucaagaaaaauagauggauuaug	ttta	514
Human-Exon 44	4	1	CAGGUaaaagcauaggaugaaga	tttt	515
Human-Exon 44	5	1	GCAGGUaaaagcauaggaugaag	tttt	516
Human-Exon 44	6	1	UGCAGGUaaaagcauaggaugaa	tttc	517

Human-Exon 44	7	-1	CAGGCGAUUUGACAGAUCUGUUGA	TTTC	518
Human-Exon 44	8	1	AGAUCUGUCAAUUCGCCUGCAGGU	tttt	519
Human-Exon 44	9	1	CAGAUCUGUCAAUUCGCCUGCAGG	tttA	520
Human-Exon 44	10	1	GCCGCCAUUUCUCAACAGAUCUGU	TTTG	521
Human-Exon 44	11	-1	AAUGGCGGCGUUUUCAUUAUGAUA	TTTA	522
Human-Exon 44	12	1	AUUAUUUAUCUUUAUAUCAUAAUG	TTTT	523
Human-Exon 44	13	-1	UGAGAAUUGGGAACAUGC UAAAUA	TTTG	524
Human-Exon 44	14	-1	GGUAAGUCUUUGAUUUGUUUUUUC	TTTC	525
Human-Exon 44	15	1	AAAUACAAUUCGAAAAACAAAU	TTTG	526
Human-Exon 44	16	1	AAGAUAAAUACAAUUCGAAAAAA	TTTG	527
Human-Exon 44	17	1	GCUGAAGAUAAAUAACAAUUCGAA	TTTT	528
Human-Exon 44	18	1	UGCUGAAGAUAAAUAACAAUUCGA	TTTT	529
Human-Exon 44	19	1	GUGCUGAAGAUAAAUAACAAUUCG	TTTT	530
Human-Exon 44	20	1	UGUGCUGAAGAUAAAUAACAAUUC	TTTC	531
Human-Exon 44	21	-1	GCACAUCUGGACUCUUUAACUUCU	TTTA	532
Human-Exon 44	22	1	UAAAGAGUCCAGAUGUGCUGAAGA	TTTA	533
Human-Exon 44	23	-1	AAGAUCAGGUUCUGAAGGGUGAUG	TTTC	534
Human-Exon 44	24	1	UUCAGAACCUGAUCUUUAAGAAGU	TTTA	535
Human-Exon 44	25	1	AAUAUAAUGAUGACAACAACAGUC	TTTT	536
Human-Exon 44	26	1	UAAUAUAAUGAUGACAACAACAGU	TTTG	537
Human-Exon 53	1	-1	UUUAUUUUUCCUUUUUAUUCUAGUU	TTTC	538
Human-Exon 53	2	1	AAAGGAAAAAUAAAUAUUAAGUAG	TTTA	539
Human-Exon 53	3	1	UUUCAACUAGAAUAAAAGGAAAAA	TTTA	540
Human-Exon 53	4	1	AUUCUUUCAACUAGAAUAAAAGGA	TTTT	541
Human-Exon 53	5	1	AAUUCUUUCAACUAGAAUAAAAGG	TTTT	542
Human-Exon 53	6	1	GAAUUCUUUCAACUAGAAUAAAAG	TTTC	543
Human-Exon 53	7	1	AUUCUGAAUUCUUUCAACUAGAAU	TTTT	544
Human-Exon 53	8	1	GAUUCUGAAUUCUUUCAACUAGAA	TTTA	545
Human-Exon 53	9	-1	CAGAACCGGAGGCAACAGUUGAAU	TTTC	546
Human-Exon 53	10	-1	GGAGGCAACAGUUGAAUGAAAUGU	TTTA	547
Human-Exon 53	11	-1	UAUACAGUAGAUGCAAUCCAAAAG	TTTT	548
Human-Exon 53	12	-1	GAUGCAAUCCAAAAGAAAUCACA	TTTC	549
Human-Exon 53	13	-1	AAUCACAGAAACCAAGGUUAGUAU	TTTG	550
Human-Exon 53	14	-1	AGGUUAGUAUCAAGAUACCUUU	TTTA	551
Human-Exon 53	15	-1	GGUUAGUAUCAAGAUACCUUUUU	TTTT	552
Human-Exon 53	16	-1	AGUAUCAAGAUACCUUUUUUAAAA	TTTA	553
Human-Exon 53	17	-1	GUAUCAAGAUACCUUUUUUAAAAU	TTTT	554
Human-Exon 46	1	-1	UGUUUGUGUCCAGUUUGCAUUA	TTTG	555
Human-Exon 46	2	1	CUGGGACACAAACAUGGCAAUUUA	TTTT	556
Human-Exon 46	3	1	ACUGGGACACAAACAUGGCAAUUU	TTTT	557
Human-Exon 46	4	1	AACUGGGACACAAACAUGGCAAUU	TTTA	558
Human-Exon 46	5	1	UAUUUGUUA AUGCAAACUGGGACA	TTTG	559

Human-Exon 46	6	-1	ACAAAUAGUUUGAGAACUAUGUUG	tttC	560
Human-Exon 46	7	-1	CAAAUAGUUUGAGAACUAUGUUGG	tttt	561
Human-Exon 46	8	-1	AAAUAGUUUGAGAACUAUGUUGGa	tttt	562
Human-Exon 46	9	-1	AUAGUUUGAGAACUAUGUUGGaaa	tttt	563
Human-Exon 46	10	-1	UAGUUUGAGAACUAUGUUGGaaaa	tttt	564
Human-Exon 46	11	-1	AGUUUGAGAACUAUGUUGGaaaaa	tttt	565
Human-Exon 46	12	1	UAGUUCUCAAAACUAUUUGUUA AUG	TTTG	566
Human-Exon 46	13	1	UAuuuuuuuuuCCAACUAUGUUCU	TTTG	567
Human-Exon 46	14	-1	CUUCUUUCUCCAGGCUAGAAGAAC	TTTT	568
Human-Exon 46	15	1	CUUCUAGCCUGGAGAAAGAAGAAU	TTTT	569
Human-Exon 46	16	1	UCUUCUAGCCUGGAGAAAGAAGAA	TTTA	570
Human-Exon 46	17	1	AUUCUUUUGUUCUUCUAGCCUGGA	TTTC	571
Human-Exon 46	18	-1	CAAAGAAUAUCUUGUCAGAAUUU	TTTG	572
Human-Exon 46	19	-1	CUGGAAAAGAGCAGCAACUAAAAG	TTTT	573
Human-Exon 46	20	-1	CAAGUCAAGGUAAUUUUU AUUUUCU	TTTG	574
Human-Exon 46	21	-1	CAAUCCCCCAGGGCCUGCUUGCA	TTTA	575
Human-Exon 46	22	1	AGGCCUGGGGGAUUUGAGAAAAU	TTTT	576
Human-Exon 46	23	1	CAGGCCUGGGGGAUUUGAGAAAA	TTTA	577
Human-Exon 46	24	1	CAAGCAGGCCUGGGGGAUUUGAG	TTTT	578
Human-Exon 46	25	1	GCAAGCAGGCCUGGGGGAUUUGA	TTTC	579
Human-Exon 46	26	1	GCAGAAAACCAAUGAUUGAAUUA	TTTT	580
Human-Exon 46	27	1	GGCAGAAAACCAAUGAUUGAAUUA	TTTT	581
Human-Exon 46	28	1	GGGCAGAAAACCAAUGAUUGAAU	TTTT	582
Human-Exon 46	29	1	UGGGCAGAAAACCAAUGAUUGAAU	TTTA	583
Human-Exon 46	30	-1	AUUAGGUUAUUC AUAGUCCUUGC	TTTA	584
Human-Exon 46	31	1	AACUAUGAAUAACCUAAUGGGCAG	TTTT	585
Human-Exon 46	32	1	GAACUAUGAAUAACCUAAUGGGCA	TTTC	586
Human-Exon 52	1	-1	UAUUUCCUGUUA AAUUGUUUUCUA	TTTA	587
Human-Exon 52	2	1	GGUUUAUAGAAAACAAUUUAACAG	TTTC	588
Human-Exon 52	3	-1	AUACAGUAACAUCUUUUUU AUUUC	TTTA	589
Human-Exon 52	4	-1	UACAGUAACAUCUUUUUU AUUUCU	TTTT	590
Human-Exon 52	5	1	AUGUUACUGUAUAAGGGUUUAUAG	TTTT	591
Human-Exon 52	6	1	GAUGUUACUGUAUAAGGGUUUAUA	TTTC	592
Human-Exon 52	7	1	CAGCCAAAACACUUUUAGAAUAA	TTTT	593
Human-Exon 52	8	1	CCAGCCAAAACACUUUUAGAAUAA	TTTT	594
Human-Exon 52	9	1	ACCAGCCAAAACACUUUUAGAAU	TTTT	595
Human-Exon 52	10	1	GACCAGCCAAAACACUUUUAGAAA	TTTA	596
Human-Exon 52	11	1	GUGAGACCAGCCAAAACACUUUA	TTTC	597
Human-Exon 52	12	-1	AAUUGUACUUUACUUUGUAUUAUG	TTTA	598
Human-Exon 52	13	-1	AUUGUACUUUACUUUGUAUUAUGU	TTTT	599
Human-Exon 52	14	1	UAAAGUACAAUUGUGAGACCAGCC	TTTT	600
Human-Exon 52	15	1	GUAAAGUACAAUUGUGAGACCAGC	TTTG	601

Human-Exon 52	16	1	GUAUUCCUUUUACAUAUACAAAG	TTTA	602
Human-Exon 52	17	1	GUUGUGUAUUCCUUUUACAUAUA	TTTG	603
Human-Exon 52	18	1	AUCCUGCAUUGUUGCCUGUAAGAA	TTTG	604
Human-Exon 52	19	1	UUCCAACUGGGGACGCCUCUGUUC	TTTG	605
Human-Exon 52	20	-1	UUGGAAGAACUCAUUAACGCUGCC	TTTG	606
Human-Exon 52	21	-1	UCAUUACGCUGCCCCAAAUUUGA	TTTT	607
Human-Exon 52	22	1	CUCUUGAUUGCUGGUCUUGUUUUU	TTTG	608
Human-Exon 52	23	-1	GUUUUUUAACAAGCAUGGGACACA	TTTG	609
Human-Exon 52	24	1	CUUUGUGUGUCCCAUGCUUGUAA	TTTT	610
Human-Exon 52	25	1	GCUUUGUGUGUCCCAUGCUUGUUA	TTTT	611
Human-Exon 52	26	1	UGC UUUGUGUGUCCCAUGCUUGUU	TTTT	612
Human-Exon 52	27	1	UUGC UUUGUGUGUCCCAUGCUUGU	TTTA	613
Human-Exon 52	28	-1	AGCAAGAUGCAUGACAAGUUUCA	TTTA	614
Human-Exon 52	29	-1	GCAAGAUGCAUGACAAGUUUCAU	TTTT	615
Human-Exon 52	30	-1	CAAGAUGCAUGACAAGUUUCAUA	TTTT	616
Human-Exon 52	31	1	GAUAUAUGAACUUAAGUUUUUAU	TTTC	617
Human-Exon 50	1	-1	AUAGAAAUCCAAUAAUUAUUCAC	TTTG	618
Human-Exon 50	2	-1	AUUAAGAUGUUC AUGAAUUAUCUU	TTTG	619
Human-Exon 50	3	-1	UAAGUAAUGUGUAUGCUUUUCUGU	TTTA	620
Human-Exon 50	4	1	AUCUUCUAACUCCUCUUUAACAG	TTTT	621
Human-Exon 50	5	1	GAUCUUCUAACUCCUCUUUAACA	TTTC	622
Human-Exon 50	6	-1	AUCUGAGCUCUGAGUGGAAGGCGG	TTTA	623
Human-Exon 50	7	-1	ACCGUUUACUUCAAGAGCUGAGGG	TTTG	624
Human-Exon 50	8	1	CUGCUUUGCCCUCAGCUCUUGAAG	TTTA	625
Human-Exon 50	9	-1	UCUCUUUGGCUCUAGCUAUUUGUU	TTTG	626
Human-Exon 50	10	-1	CUCUUUGGCUCUAGCUAUUUGUUC	TTTT	627
Human-Exon 50	11	1	CACUUUUGAACAAAUAGCUAGAGC	TTTG	628
Human-Exon 50	12	1	UCACUUCAUAGUUGCACUUUUGAA	TTTG	629
Human-Exon 50	13	-1	AUGAAGUGAUGACUGGGUGAGAGA	TTTC	630
Human-Exon 50	14	-1	UGAAGUGAUGACUGGGUGAGAGAG	TTTT	631
Human-Exon 43	1	1	AAGAGAAA <u>Auauauauauaua</u>	TTTG	632
Human-Exon 43	2	1	GAAUUAGCUGUCUAUAGAAAGAGA	tTTT	633
Human-Exon 43	3	1	UGAAUUAGCUGUCUAUAGAAAGAG	TTTT	634
Human-Exon 43	4	-1	AGCUAAUUCAUUUUUUUACUGUUU	TTTA	635
Human-Exon 43	5	1	AUGAAUUAGCUGUCUAUAGAAAGA	TTTC	636
Human-Exon 43	6	-1	GCUAAUUCAUUUUUUUACUGUUUU	TTTT	637
Human-Exon 43	7	1	AAAAAAUUGAAUUAGCUGUCUAUA	TTTC	638
Human-Exon 43	8	-1	UUAAAAUUUUUAUUAUACAGAAUA	TTTA	639
Human-Exon 43	9	-1	UAAAAUUUUUAUUAUACAGAAUAU	TTTT	640
Human-Exon 43	10	1	AUAUAAAAUUUUAAAACAGUAAA	TTTT	641
Human-Exon 43	11	1	AAUAUAAAAUUUUAAAACAGUAA	TTTT	642
Human-Exon 43	12	1	UAAUAUAAAAUUUUAAAACAGUA	TTTT	643

Human-Exon 43	13	1	GUAAUAUAAAAAUUUUAAAACAGU	TTTT	644
Human-Exon 43	14	1	UGUAAUAUAAAAAUUUUAAAACAG	TTTA	645
Human-Exon 43	15	1	UAUAUUCUGUAAUAUAAAAAUUUU	TTTT	646
Human-Exon 43	16	1	UUAUAUUCUGUAAUAUAAAAAUUU	TTTA	647
Human-Exon 43	17	-1	CAGAAUAUAAAAGAUAGUCUACAA	TTTG	648
Human-Exon 43	18	1	CUAUCUUUUUAUAUUCUGUAAUAUA	TTTT	649
Human-Exon 43	19	1	ACUAUCUUUUUAUAUUCUGUAAUAU	TTTT	650
Human-Exon 43	20	1	GACUAUCUUUUUAUAUUCUGUAAUA	TTTA	651
Human-Exon 43	21	-1	CAUAGCAAGAAGACAGCAGCAUUG	TTTG	652
Human-Exon 43	22	1	CAUUUUGUUAACUUUUUCCCAUUG	TTTC	653
Human-Exon 43	23	-1	CAUAUAUUUUUCUUGAUACUUGCA	TTTC	654
Human-Exon 43	24	1	AAAUCAUUUCUGCAAGUAUCAAGA	TTTT	655
Human-Exon 43	25	1	CAAUCAUUUCUGCAAGUAUCAAG	TTTT	656
Human-Exon 43	26	1	ACAAUCAUUUCUGCAAGUAUCAAA	TTTC	657
Human-Exon 43	27	1	AUAAAUUCUACAGUUCCCGUAAAA	TTTG	658
Human-Exon 43	28	-1	GAAUUUAUUUCAGUACCCUCCAUG	TTTC	659
Human-Exon 43	29	-1	AAUUUAUUUCAGUACCCUCCAUGG	TTTT	660
Human-Exon 43	30	1	UGAAAUAAAUUCUACAGUUCCCGU	TTTT	661
Human-Exon 43	31	-1	AUUUAUUUCAGUACCCUCCAUGGA	TTTT	662
Human-Exon 43	32	1	CUGAAAUAAAUUCUACAGUUCCCU	TTTC	663
Human-Exon 43	33	-1	UUUAUUUCAGUACCCUCCAUGGAA	TTTT	664
Human-Exon 43	34	-1	UACCCUCCAUGGAAAAAAGACAGG	TTTC	665
Human-Exon 43	35	-1	ACCCUCCAUGGAAAAAAGACAGGG	TTTT	666
Human-Exon 43	36	-1	CCCUCUCCAUGGAAAAAAGACAGGGA	TTTT	667
Human-Exon 43	37	1	UUUUUCCAUGGAGGGUACUGAAA	TTTA	668
Human-Exon 43	38	1	UGUCUUUUUCCAUGGAGGGUACU	TTTC	669
Human-Exon 6	1	1	CCUUGAGCAAGAACCAUGCAAACU	TTTA	670
Human-Exon 6	2	-1	UGCUC AAGGAUGCAUUUUCUUAU	TTTC	671
Human-Exon 6	3	-1	GCUCAAGGAUGCAUUUUCUUAUG	TTTT	672
Human-Exon 6	4	1	UGCAUUCUUGAGCAAGAACCAUG	TTTG	673
Human-Exon 6	5	-1	GAAAAUUUAUUUCCACAUGUAGGU	TTTG	674
Human-Exon 6	6	-1	AAAAUUUAUUUCCACAUGUAGGUC	TTTT	675
Human-Exon 6	7	-1	AAAUUUUAUUUCCACAUGUAGGUCA	TTTT	676
Human-Exon 6	8	1	CAUGUGGAAAUAAAUUUUCAUAA	TTTT	677
Human-Exon 6	9	1	ACAUGUGGAAAUAAAUUUUCAUAA	TTTC	678
Human-Exon 6	10	-1	CCACAUGUAGGUCAAAAUGUAAU	TTTC	679
Human-Exon 6	11	-1	CACAUGUAGGUCAAAAUGUAAUG	TTTT	680
Human-Exon 6	12	-1	ACAUGUAGGUCAAAAUGUAAUGA	TTTT	681
Human-Exon 6	13	1	ACAUUUUUGACCUACAUGUGGAAA	TTTA	682
Human-Exon 6	14	1	CAUUACAUUUUUGACCUACAUGUG	TTTC	683
Human-Exon 6	15	-1	AAAAUAUCAUGGCUGGAUUGCAA	TTTG	684
Human-Exon 6	16	-1	GCUGGAUUGCAACAAACCAACAGU	TTTC	685

Human-Exon 6	17	-1	CUGGAUUGCAACAAACCAACAGUG	TTTT	686
Human-Exon 6	18	1	CCUAUGACUAUGGAUGAGAGCAUU	TTTG	687
Human-Exon 6	19	-1	UAGGUAAGAAGAUUACUGAGACAU	TTTA	688
Human-Exon 6	20	-1	AUUACUGAGACAUUAAAUAACUUG	TTTA	689
Human-Exon 6	21	-1	UUACUGAGACAUUAAAUAACUUGU	TTTT	690
Human-Exon 6	22	1	GGGGAAAAUAUGUCAUCAGAGUC	TTTA	691
Human-Exon 6	23	1	CAUGAUCUGGAACCAUACUGGGGA	TTTT	692
Human-Exon 6	24	1	ACAUGAUCUGGAACCAUACUGGGG	TTTT	693
Human-Exon 6	25	1	GACAUGAUCUGGAACCAUACUGGG	TTTC	694
Human-Exon 7	1	1	uacacacauacacaAAGACAAUA	TTTA	695
Human-Exon 7	2	1	uacacauacacacauacacaAGA	TTTG	696
Human-Exon 7	3	1	aacacauacacauacacacauaca	TTtg	697
Human-Exon 7	4	1	AUCCAGUCAAAUAGGUCUGGCCU	ttTT	698
Human-Exon 7	5	1	UAUCCAGUCAAAUAGGUCUGGCC	tTTA	699
Human-Exon 7	6	1	GCUGGCAAACCACACUAUUCAGU	TTTG	700
Human-Exon 7	7	1	AGUCGUUGUGUGGCUGACUGCUGG	TTTG	701
Human-Exon 7	8	-1	CGCCAGAUAUCAAUAGGCAUAGA	TTTC	702
Human-Exon 7	9	-1	AAACUACUCGAUCCUGAAGGUUGG	TTTA	703
Human-Exon 7	10	1	CAUACUAAAAGCAGUGGUAGUCCA	TTTC	704
Human-Exon 7	11	1	GAAAACAUUAAACUCUACCAUACU	TTTT	705
Human-Exon 7	12	1	UGAAAACAUUAAACUCUACCAUAC	TTTA	706
Human-Exon 8	1	-1	UUGUUCAUUAUCCUUUUAGAGUCU	TTTG	707
Human-Exon 8	2	1	AAAGGAUAAUGAACAAAUCAAAGU	TTTA	708
Human-Exon 8	3	-1	UAUCCUUUUAGAGUCUCAAAUAUA	TTTC	709
Human-Exon 8	4	1	ACUCUAAAAGGAUAAUGAACAAAU	TTTG	710
Human-Exon 8	5	-1	UUUUAGAGUCUCAAAUAUAGAAAC	TTTG	711
Human-Exon 8	6	-1	UUUAGAGUCUCAAAUAUAGAAACC	TTTT	712
Human-Exon 8	7	-1	UUAGAGUCUCAAAUAUAGAAACCA	TTTT	713
Human-Exon 8	8	1	UUGAGACUCUAAAAGGAUAAUGAA	TTTG	714
Human-Exon 8	9	1	UUUGGUUUCUAUAUUUGAGACUCU	TTTT	715
Human-Exon 8	10	1	UUUUGGUUUCUAUAUUUGAGACUC	TTTA	716
Human-Exon 8	11	-1	AGCAUUGAAGCCAUCCAGGAAGUG	TTTC	717
Human-Exon 8	12	1	GCUUCA AUGCUCACUUGUUGAGGC	TTTT	718
Human-Exon 8	13	1	GGCUUCA AUGCUCACUUGUUGAGG	TTTG	719
Human-Exon 8	14	-1	AGUGGAA AUGUUGCCAAGGCCACC	TTTA	720
Human-Exon 8	15	-1	GUUGCCAAGGCCACC UAAAGUGAC	TTTA	721
Human-Exon 8	16	-1	GAAGAACAUUUUCAGUUA CAUCAU	TTTG	722
Human-Exon 8	17	-1	AUCAA AUGCACUAUUCUCAACAGG	TTTA	723
Human-Exon 8	18	1	AUAGUGCAUUGAUGAUGUAACUG	TTTT	724
Human-Exon 8	19	1	AAUAGUGCAUUGAUGAUGUAACU	TTTC	725
Human-Exon 8	20	-1	ACUAUUCUCAACAGGUAAAGUGUG	TTTA	726
Human-Exon 8	21	1	UACCUAAAA AUGCAUAUAAAACAG	TTTT	727

Human-Exon 8	22	1	AUACCUAAAAAUGCAUAUAAAACA	TTTC	728
Human-Exon 8	23	1	CACGUAAUACCUAAAAAUGCAUAU	TTTT	729
Human-Exon 8	24	1	GCACGUAAUACCUAAAAAUGCAUA	TTTA	730
Human-Exon 8	25	1	auauauauGUGCACGUAAUACCUA	TTTT	731
Human-Exon 8	26	1	uauauauauGUGCACGUAAUACCU	TTTT	732
Human-Exon 8	27	1	auauauauauGUGCACGUAAUACC	TTTA	733
Human-Exon 55	1	-1	CUGCACAAUAUUAUAGUUGUUGCU	TTTA	734
Human-Exon 55	2	1	AUAAAAAGAGAAAGAUGGAGGAAC	TTTA	735
Human-Exon 55	3	1	CACCUAGUGAACUCCAUAAAAAGA	TTTC	736
Human-Exon 55	4	1	AUGGUGCACCUAGUGAACUCCAUA	TTTT	737
Human-Exon 55	5	1	AAUGGUGCACCUAGUGAACUCCAUA	TTTT	738
Human-Exon 55	6	1	GAAUGGUGCACCUAGUGAACUCCA	TTTA	739
Human-Exon 55	7	1	GACCAAUGUUCAGAUGCAAUUAU	TTTA	740
Human-Exon 55	8	1	UCGCUCACUCACCCUGCAAAGGAC	TTTG	741
Human-Exon 55	9	-1	AGUGAGCGAGAGGCUGCUUUGGAA	TTTC	742
Human-Exon 55	10	1	GCAGCCUCUCGCUCACUCACCCUG	TTTG	743
Human-Exon 55	11	1	UUGCAGUAAUCUAUGAGUUUCUUC	TTTG	744
Human-Exon 55	12	-1	CUGCAACAGUUCUUUUUGGACCUG	TTTC	745
Human-Exon 55	13	-1	UGCAACAGUUCUUUUUGGACCUGG	TTTT	746
Human-Exon 55	14	-1	UUUCUUGCCUGGCUUACAGAAGCU	TTTC	747
Human-Exon 55	15	1	UUUCAGCUUCUGUAAGCCAGGCAA	TTTC	748
Human-Exon 55	16	-1	GUCCUACAGGAUGCUACCCGUAAG	TTTC	749
Human-Exon 55	17	-1	GGCUCCUAGAAGACUCCAAGGGAG	TTTA	750
Human-Exon 55	18	-1	GCUCCUAGAAGACUCCAAGGGAGU	TTTT	751
Human-Exon 55	19	-1	CUCCAAGGGAGUAAAAGAGCUGAU	TTTC	752
Human-Exon 55	20	1	UGGAUCCACAAGAGUGCUAAAGCG	TTTC	753
Human-Exon 55	21	1	GUUCAAUUGGAUCCACAAGAGUGC	TTTA	754
Human-Exon 55	22	-1	UACUUGUAACUGACAAGCCAGGGA	TTTG	755
Human-Exon 55	23	-1	ACUUGUAACUGACAAGCCAGGGAC	TTTT	756
Human-Exon 55	24	-1	GUAACUGACAAGCCAGGGACAAAA	TTTG	757
Human-Exon 55	25	-1	UAACUGACAAGCCAGGGACAAAAC	TTTT	758
Human-Exon 55	26	1	UCCUGGCUUGUCAGUUACAAGUA	TTTG	759
Human-G1-exon51		1	CAGAGUAACAGUCUGAGUAGGAGc	TTTA	760
Human-G2-exon51		1	uacuuuguuuagcauacauggua	TTTC	761
Human-G3-exon51		-1	uggcucaaauguuacucucaau	TTTA	762
mouse-Exon23-G1		1	CUUUCAAagaacuuugcagagccu	TTTG	763
mouse-Exon23-G2		1	guugaaGCCAUUUUGUUGCUCUUU	TTTG	764
mouse-Exon23-G3		1	guugaaGCCAUUUUAUUGCUCUUU	TTTG	765

mouse-Exon23-G4		-1	uuuugagGCUCUGCAAAGUUCUUU	TTTC	766
mouse-Exon23-G5		-1	aguuaauaaugcauagauauucag	TTTA	767
mouse-Exon23-G6		-1	uauaauaugcccuguaauuaaua	TTTC	768
mouse-Exon23-G7		1	uaaaggccaaaccucggcuuaccU	TTTC	769
mouse-Exon23-G8		1	ucaauaucuuugaaggacucuggg	TTTA	770

* In this table, upper case letters represent sgRNA nucleotides that align to the exon sequence of the gene. Lower case letters represent sgRNA nucleotides that align to the intron sequence of the gene.

5 VI. Examples

The following examples are included to demonstrate preferred embodiments of the disclosure. It should be appreciated by those of skill in the art that the techniques disclosed in the examples which follow represent techniques discovered by the inventor to function well in the practice of the disclosure, and thus can be considered to constitute preferred modes for its practice. However, those of skill in the art should, in light of the present disclosure, appreciate that many changes can be made in the specific embodiments which are disclosed and still obtain a like or similar result without departing from the spirit and scope of the disclosure.

15 EXAMPLE 1 - Materials and Methods

Generation of pLbCpf1-2A-GFP and pAsCpf1-2A-GFP plasmids. Human codon-optimized LbCpf1 and AsCpf1 were PCR amplified from pY016 plasmid (Zetsche *et al.*, 2015) (pcDNA3.1-hLbCpf1), a gift from Feng Zhang (Addgene plasmid # 69988) and pY010 plasmid (Zetsche *et al.*, 2015) (pcDNA3.1-hAsCpf1), a gift from Feng Zhang (Addgene plasmid # 69982), respectively. Cpf1 cDNA and T2A-GFP DNA fragment were cloned into the backbone of the pSpCas9(BB)-2A-GFP (PX458) plasmid (Ran *et al.*, 2015), a gift from Feng Zhang (Addgene plasmid # 48138) that was cut with AgeI/EcoRI to remove SpCas9(BB)-2A-GFP. In-Fusion HD cloning kit (Takara Bio) was used. Cpf1 guide RNAs (gRNAs) targeting the human *DMD* or the mouse *Dmd* locus were sub-cloned into a newly generated pLbCpf1-2A-GFP plasmid and pAsCpf1-2A-GFP plasmid using BbsI digestion and T4 ligation. Detailed primer sequences can be found in Table C, genomic target sequences can be found in Table D, and gRNA sequences can be found in Table E.

Human iPSC maintenance, nucleofection and differentiation. Human iPSCs were cultured in mTeSRTM1 media (STEMCELL Technologies) and passaged approximately every 4 days (1:18 split ratio). One hour before nucleofection, iPSCs were treated with 10 μ M ROCK inhibitor (Y-27632) and dissociated into single cells using Accutase (Innovative Cell Technologies, Inc.). 1 x 10⁶ iPSC cells were mixed with 5 μ g of pLbCpf1-2A-GFP or pAsCpf1-2A-GFP plasmid and nucleofected using the P3 Primary Cell 4D-Nucleofector X kit (Lonza) according to manufacturer's protocol. After nucleofection, iPSCs were cultured in mTeSRTM1 media supplemented with 10 μ M ROCK inhibitor, penicillin-streptomycin (1:100) (ThermoFisher Scientific) and 100 μ g/ml Primosin (InvivoGen). Three days post-nucleofection, GFP(+) and (-) cells were sorted by FACS and subjected to T7E1 assay. Single clones derived from GFP(+). iPSCs were picked and sequenced. iPSCs were induced to differentiate into cardiomyocytes, as previously described (Burridge *et al.* 2014).

Genomic DNA isolation. Genomic DNA of mouse 10T1/2 fibroblasts and human iPSCs was isolated using Quick-gDNA MiniPrep kit (Zymo Research) according to manufacturer's protocol.

RT-PCR. RNA was isolated using TRIzol (ThermoFisher Scientific), according to manufacturer's protocol. cDNA was synthesized using iScript Reverse Transcription Supermix (Bio-Rad Laboratories) according to manufacturer's protocol. RT-PCR was performed using primers flanking *DMD* exon 47 and 52:

forward: 5'-CCCAGAAGAGCAAGATAAACTTGAA-3' (SEQ ID NO: 1);

reverse: 5'-CTCTGTTCCAAATCCTGCTTGT-3' (SEQ ID NO: 2)

RT-PCR products amplified from WT cardiomyocytes, uncorrected cardiomyocytes and exon 51 skipped cardiomyocytes were 717 bps, 320 bps and 87 bps, respectively.

Dystrophin Western blot analysis. Western blot analysis were performed as previously described (Long *et al.*, 2014) using rabbit anti-dystrophin antibody (Abcam, ab15277) and mouse anti-cardiac myosin heavy chain antibody (Abcam, ab50967).

Dystrophin immunocytochemistry and immunohistochemistry. iPSC-derived cardiomyocytes fixed with acetone, blocked with serum cocktail (2% normal horse serum/2% normal donkey serum/0.2% bovine serum albumin (BSA)/PBS), and incubated with dystrophin antibody (MANDYS8, 1:800, Sigma-Aldrich) and troponin-I antibody (H170, 1:200, Santa Cruz Biotechnology) in 0.2% BSA/PBS. Following overnight incubation at 4°C, they were incubated with secondary antibodies (biotinylated horse anti-mouse IgG, 1:200, Vector Laboratories, fluorescein-conjugated donkey anti-rabbit IgG, 1:50, Jackson Immunoresearch) for one-hour. Nuclei were counterstained with Hoechst 33342 (Molecular Probes).

Immunohistochemistry of skeletal muscle was performed as previously described (Long *et al.*, 2014) using dystrophin antibody (MANDYS8, 1:800, Sigma-Aldrich). Nuclei were counterstained with propidium iodide (Molecular Probes).

Mitochondrial DNA copy number quantification. Genomic and mitochondrial DNA were isolated using Trizol, followed by back extraction as previously described (Zechner *et al.*, 2010). KAPA SYBR FAST qPCR kit (Kapa Biosystems) was used to perform real-time PCR to quantitatively determine mitochondrial DNA copy number. Human mitochondrial ND1 gene was amplified using primers (forward: 5'- CGCCACATCTACCATCACCTC -3' (SEQ ID NO: 3); reverse: 5'- CGGCTAGGCTAGAGGTGGCTA -3' (SEQ ID NO: 4)). Human genomic LPL gene was amplified using primers (forward: 5'- GAGTATGCAGAAGCCCCGAGTC -3' (SEQ ID NO: 5); reverse: 5'- TCAACATGCCCAACTGGTTTCTGG -3' (SEQ ID NO: 6)). mtDNA copy number per diploid genome was calculated using formula:

$$\Delta CT = (\text{mtND1 CT} - \text{LPL CT})$$

$$\text{mtDNA copy number per diploid genome} = 2 \times 2^{-\Delta CT}$$

Cellular respiration rates. Oxygen consumption rates (OCR) were determined in human iPSC-derived cardiomyocytes using the XF24 Extracellular Flux Analyzer (Seahorse Bioscience) following the manufacturer's protocol as previously described (Baskin *et al.*, 2014).

In vitro transcription of LbCpf1 mRNA and gRNA. Human codon-optimized LbCpf1 was PCR amplified from pLbCpf1-2A-GFP to include the T7 promoter sequence (Table S1). The PCR product was transcribed using mMESSAGE mMACHINE T7 transcription kit (ThermoFisher Scientific) according to manufacturer's protocol. Synthesized LbCpf1 mRNA were poly-A tailed with *E. coli* Poly(A) Polymerase (New England Biolabs) and purified using NucAway spin columns (ThermoFisher Scientific).

The template for LbCpf1 gRNA *in vitro* transcription was PCR amplified from pLbCpf1-2A-GFP plasmid and purified using Wizard SV gel and PCR clean-up system (Promega). The LbCpf1 gRNA was synthesized using MEGAshortscript T7 transcription kit (ThermoFisher Scientific) according to manufacturer's protocol. Synthesized LbCpf1 gRNA were purified using NucAway spin columns (ThermoFisher Scientific).

Single-stranded oligodeoxynucleotide (ssODN). ssODN was used as HDR template and synthesized by Integrated DNA Technologies as 4nM Ultramer Oligonucleotides. ssODN was mixed with LbCpf1 mRNA and gRNA directly without purification. The sequence of ssODN is:

5'-

TGATATGAATGAACTCATCAAATATGCGTGTTAGTGTAATGAACTTCTATTTA
ATTTTGAGGCTCTGCAAAGTTCTTTAAAGGAGCAGCAGAATGGCTTCAACTATCT
GAGTGACACTGTGAAGGAGATGGCCAAGAAAGCACCTTCAGAAATATGCCAGAA
5 ATATCTGTCAGAATTT-3' (SEQ ID NO: 7)

CRISPR-Cpf1-mediated genome editing by one-cell embryo injection. All animal procedures were approved by the Institutional Animal Care and Use Committee at the University of Texas Southwestern Medical Center. Injection procedures were performed as described previously (Long *et al.*, 2014). The only modification was replacing Cas9 mRNA and Cas9 sgRNAs with LbCpf1 mRNA and LbCpf1 gRNAs.

PCR amplification of genomic DNA, T7E1 assay, and TseI RFLP analysis. These methods were performed as previously published (Long *et al.*, 2014).

Statistical analysis. Statistical analysis was assessed by two-tailed Student's t-test. Data are shown as mean $\pm$ SEM. A $P < 0.05$ value was considered statistically significant.

EXAMPLE 2 - Results

Correction of DMD iPSC-derived cardiomyocytes by Cpf1-mediated genome editing. Exon deletions preceding exon 51 of the human *DMD* gene, which disrupt the open reading frame (ORF) by juxtaposing out of frame exons, represent the most common type of human DMD mutation (Aartsma-Rus *et al.*, 2009). Skipping of exon 51 can, in principle, restore the DMD ORF in 13% of DMD patients with exon deletions (Cirak *et al.*, 2011). To test the potential of Cpf1 to correct this type of "hot-spot" mutation, the inventors used DMD fibroblast-derived iPSCs (Riken HPS0164, abbreviated as Riken51), which harbor a deletion of exons 48 to 50, introducing a premature termination codon within exon 51 (FIG. 1A).

The splice acceptor region is generally T/C-rich (Padgett, 2012), which creates an ideal PAM sequence for genome editing by Cpf1 endonuclease (FIG. 1B). To rescue dystrophin expression in Riken51 iPSCs, the inventors used a Cpf1 gRNA to target exon 51, introducing small insertions and deletions (INDELs) in exon 51 by NHEJ and subsequently reframing the dystrophin ORF, theoretically, in one-third of corrected genes, a process inventors refer to as "reframing" (FIG. 1A). They also compared two Cpf1 orthologs, LbCpf1 (from *Lachnospiraceae bacterium sp. ND2006*; UniProt Accession No. A0A182DWE3; SEQ ID No. 443) and AsCpf1 (from *Acidaminococcus sp. BV3L6*; SEQ ID No. 442), which use the same PAM sequences for genome cleavage.

Cpf1 cleavage was targeted to the T-rich splice acceptor site of exon 51 using a guide RNA (designated g1) (FIG. 1C), which was cloned into plasmids pLbCpf1-2A-GFP and pAsCpf1-2A-GFP (FIG. 1D). These plasmids express human codon optimized LbCpf1 or AsCpf1, plus GFP; enabling fluorescence activated cell sorting (FACS) of Cpf1-expressing cells (FIG. 1D). Initially, inventors evaluated the cleavage efficiency of Cpf1-editing with g1 in human 293T cells. Both LbCpf1 and AsCpf1 efficiently induced DNA cleavage with g1, as detected using a T7E1 assay that recognizes and cleaves non-perfectly matched DNA (FIG. 1E).

Next, inventors used LbCpf1 and AsCpf1 with g1 to edit Riken51 iPSCs, and by the T7E1 assay the inventors observed genome cleavage at *DMD* exon 51 (FIG. 1E). Genomic PCR products from the Cpf1-edited *DMD* exon 51 were cloned and sequenced (FIG. 6A). They observed INDELs near the exon 51 splice acceptor site in both LbCpf1- and AsCpf1-edited Riken51 iPSCs (FIG. 6A). Single clones from a mixture of reframed Riken51 iPSCs were picked and expanded and the edited genomic region was sequenced. Out of 12 clones, inventors observed four clones with reframed *DMD* exon 51, which restored the ORF (FIG. 6B).

Restoration of dystrophin expression in DMD iPSC-derived cardiomyocytes after Cpf1-mediated reframing. Riken51 iPSCs edited by CRISPR-Cpf1 using the reframing strategy were induced to differentiate into cardiomyocytes (Burridge *et al.*, 2014) (FIG. 2A). Cardiomyocytes with the reframed *DMD* gene were identified by RT-PCR using a forward primer targeting exon 47 and a reverse primer targeting exon 52 and PCR products were sequenced (FIGS. 2B-C). Uncorrected iPSC-derived cardiomyocytes have a premature termination codon following the first 8 amino acids encoded by exon 51, which creates a premature stop codon (FIG. 2C). Cardiomyocytes differentiated from Cpf1-edited Riken51 iPSCs showed restoration of the *DMD* ORF as seen by sequencing of the RT-PCR products from amplification of exons 47 to 52 (FIG. 2C). The inventors also confirmed restoration of dystrophin protein expression by Western blot analysis and immunocytochemistry using dystrophin antibody (FIGS. 2D-E). Surprisingly, even without clonal selection and expansion, cardiomyocytes differentiated from Cpf1-edited iPSC mixtures showed levels of dystrophin protein comparable to WT cardiomyocytes (FIG. 2D).

From mixtures of LbCpf1-edited Riken51 iPSCs, the inventors picked two clones (clone #2 and #5) with in-frame INDELs of different sizes and differentiated the clones into cardiomyocytes. Clone #2 had an 8 bp deletion at the 5'-end of exon 51, together with an endogenous deletion of exons 48-50. The total 405 bp deletion restored the *DMD* ORF and allowed for the production of a truncated dystrophin protein with a 135 amino acid deletion.

Clone #5 had a 17 bp deletion in exon 51 and produced dystrophin protein with a 138 amino acid deletion. Although there is high efficiency of cleavage by Cpf1, the amount of DNA inserted or deleted at the cleavage site varies. Additionally, INDELs can generate extra codons at the edited locus, causing changes of the ORF. The dystrophin protein expressed by clone #2
5 cardiomyocytes generated an additional four amino acids (Leu-Leu-Leu-Arg) between exon 47 and exon 51, whereas dystrophin protein expressed by clone #5 cardiomyocytes generated only one additional amino acid (Leu). From both clones #2 and #5, the inventors observed restored dystrophin protein by Western blot analysis and immunocytochemistry (FIGS. 2F-G). Due to the large size of dystrophin, the internally-deleted forms migrated similarly to WT dystrophin
10 on SDS-PAGE.

The inventors also performed functional analysis of DMD iPSC-derived cardiomyocytes by measuring mitochondrial DNA copy number and cellular respiration rates. Uncorrected DMD iPSC-derived cardiomyocytes had significantly fewer mitochondria than the LbCpf1-corrected cardiomyocytes (FIG. 2H). After LbCpf1-mediated reframing, both
15 corrected clones restored mitochondrial number to a level comparable to that of WT cardiomyocytes (FIG. 2H). Clone #2 iPSC-derived cardiomyocytes also showed an increase in oxygen consumption rate (OCR) compared to uncorrected iPSC-derived cardiomyocytes at baseline (FIG. 2I). OCR was inhibited by oligomycin in all iPSC-derived cardiomyocytes, and treatment with the uncoupling agent FCCP enhanced OCR. Finally, treatment with rotenone
20 and antimycin A further inhibited OCR in all cardiomyocytes. These results demonstrate that Cpf1-mediated *DMD* correction improved respiratory capacity of mitochondria in corrected iPSC-cardiomyocytes. Our findings show that Cpf1-mediated reframing is a highly efficient strategy to rescue DMD phenotypes in human cardiomyocytes.

Restoration of dystrophin expression in DMD iPSC-derived cardiomyocytes by Cpf1-mediated exon skipping. In contrast to the single gRNA-mediated reframing method,
25 which introduces small INDELs, exon skipping uses two gRNAs to disrupt splice sites and generates a large deletion (FIG. 3A). As an independent strategy to restore dystrophin expression in the Riken51 iPSCs, the inventors designed two LbCpf1 gRNAs (g2 and g3) that target the 3'-end of intron 50 and tested the cleavage efficiency in human 293T cells. T7E1
30 assay showed that g2 had higher cleavage efficiency within intron 50 compared to g3 (FIG. 3B). Therefore, the inventors co-delivered LbCpf1, g2 and g1 (g1 targets the 5' region of exon 51) into Riken51 iPSCs with the aim of disrupting the splice acceptor site of exon 51. Genomic PCR showed a lower band in LbCpf1-edited iPSCs (FIG. 3C) and sequencing data confirmed the presence of a deletion of ~200 bp between intron 50 and exon 51, which disrupted the

conserved splice acceptor site (FIG. 3D).

Riken51 iPSCs edited by the exon skipping strategy with g1 and g2 were differentiated into cardiomyocytes. Cells harboring the edited *DMD* allele were identified by RT-PCR using a forward primer targeting exon 47 and a reverse primer targeting exon 52; showing deletion of the exon 51 splice acceptor site which allows skipping of exon 51 (FIG. 3E). Sequencing of the RT-PCR products confirmed that exon 47 was spliced to exon 52, which restored the *DMD* ORF (FIG. 3F). Western blot analysis and immunocytochemistry confirmed the restoration of dystrophin protein expression in a mixture of LbCpf1-edited cardiomyocytes with g1 and g2 (FIGS. 3G-H). Thus, Cpf1-editing by the exon skipping strategy is highly efficient in rescuing the *DMD* phenotype in human cardiomyocytes.

Restoration of dystrophin in *mdx* mice by Cpf1-mediated correction. To further evaluate the potential of Cpf1-mediated *Dmd* correction *in vivo*, the inventors used LbCpf1 to permanently correct the mutation in germline of *mdx* mice by HDR-mediated correction or NHEJ-mediated reframing. *mdx* mice carry a nonsense mutation in exon 23 of the *Dmd* gene, due to a C to T transition (FIG. 4A). Three gRNAs (g1, g2 and g3) that target exon 23 were screened and tested in mouse 10T1/2 fibroblasts for cleavage efficiency (FIG. 4B). The T7E1 assay revealed that LbCpf1 and AsCpf1 had different cleavage efficiencies at *Dmd* exon 23 (FIG. 4C). Based on sequencing results, LbCpf1-mediated genome editing using g2 generated a greater occurrence of INDELs in mouse fibroblasts compared to g3 (FIG. 6C).

LbCpf1-editing with g2 recognizes a PAM sequence 9 bps upstream of the mutation site and creates a staggered double-stranded DNA cut 8 bps downstream of the mutation site (FIG. 4D). To obtain HDR genome editing, the inventors used a 180 bp single-stranded oligodeoxynucleotide (ssODN) in combination with LbCpf1 and g2 since it has been shown that ssODNs are more efficient in introducing genomic modification than double-stranded donor plasmids (Wu *et al.*, 2013; Long *et al.*, 2014). They generated a ssODN containing 90 bp of homology sequence flanking the cleavage site, including, four silent mutations and a TseI restriction site to facilitate genotyping as previously described (Long *et al.*, 2014). This ssODN was designed to be used with LbCpf1 and g2 to correct the C to T mutation within *Dmd* exon 23 and to restore dystrophin in *mdx* mice by HDR.

Correction of muscular dystrophy in *mdx* mice by LbCpf1-mediated HDR. *mdx* zygotes were co-injected with *in vitro* transcribed LbCpf1 mRNA, *in vitro* transcribed g2 gRNA and 180 bp ssODN and re-implanted into pseudo-pregnant females (FIG. 5A). Three litters of LbCpf1-edited *mdx* mice were analyzed by T7E1 assay and TseI RFLP (restriction fragment length polymorphism) (FIGS. 5B-C). Out of 24 pups born, 12 were T7E1 positive

and 5 carried corrected alleles (*mdx* C1-C5), as detected by TseI RFLP and sequencing (FIGS. 5C-D). Skeletal muscles (tibialis anterior and gastrocnemius-plantaris) from WT, *mdx* and LbCpf1-corrected *mdx-C* mice were analyzed at 4 weeks of age. Hematoxylin and eosin (H&E) staining of muscle showed fibrosis and inflammatory infiltration in *mdx* muscle, whereas LbCpf1-corrected (*mdx-C*) muscle displayed normal muscle morphology and no signs of a dystrophic phenotype (Fig. 5E and FIGS. 7A-B). Immunohistochemistry showed absence of dystrophin-positive fibers in muscle sections of *mdx* mice, whereas *mdx-C* muscle corrected by LbCpf1-mediated HDR showed dystrophin protein expression in a majority of muscle fibers (FIGS. 5F and FIGS. 7A-B). These findings show that LbCpf1-mediated editing of germline DNA can effectively prevent muscular dystrophy in mice.

EXAMPLE 3 - Discussion

In this study, the inventors show that the newly discovered CRISPR-Cpf1 nuclease can efficiently correct DMD mutations in patient-derived iPSCs and *mdx* mice, allowing for restoration of dystrophin expression. Lack of dystrophin in DMD has been shown to disrupt integrity of the sarcolemma, causing mitochondria dysfunction and oxidative stress (Millay *et al.*, 2008; Mourkioti *et al.*, 2013). They found increased mitochondrial DNA and higher oxygen consumption rates in LbCpf1-corrected iPSC-derived cardiomyocytes compared to uncorrected DMD iPSC-derived cardiomyocytes. Metabolic abnormalities of human DMD iPSC-derived cardiomyocytes were also rescued by Cpf1-mediated genomic editing. The inventors' findings also demonstrated the robustness and efficiency of Cpf1 in mouse genome editing. By using HDR-mediated correction, the ORF of the mouse *Dmd* gene was completely restored and pathophysiological hallmarks of the dystrophic phenotype such as fibrosis and inflammatory infiltration were also rescued.

Two different strategies - "reframing" and "exon skipping" - were applied to restore the ORF of the *DMD* gene using LbCpf1-mediated genome editing. Reframing creates small INDELs and restores the ORF by placing out-of-frame codons in-frame. Only one gRNA is required for reframing. Although the inventors did not observe any differences in subcellular localization between WT dystrophin protein and reframed dystrophin protein by immunocytochemistry, they observed differences in dystrophin expression level, mitochondrial DNA quantity, and oxygen consumption rate in separate edited clones, suggesting that reframed dystrophin may not be structurally or functionally identical to WT dystrophin.

Various issues should be considered with respect to the use of one or two gRNAs with Cpf1-editing. Here, the inventors show that two gRNAs are more effective than one gRNA for disruption of the splice acceptor site compared to reframing. When using two gRNAs, Cpf1-editing excises the intervening region and not only removes the splice acceptor site but can be designed to remove deleterious “AG” nucleotides, eliminating the possibility of generating a pseudo-splice acceptor site. However, with two gRNAs there is the necessity that both gRNAs cleave simultaneously, which may not occur. If only one of the two gRNAs cleaves, the desired deletion will not be generated. However, there remains the possibility that cleavage with one of the two gRNAs will generate INDELS at the targeted exon region, reframing the ORF, since in theory, one third of the INDELS will be in-frame. Using one gRNA to disrupt the splice acceptor site seems more efficient because it eliminates the need for two simultaneous cuts to occur. However, there is uncertainty with respect to the length of the INDEL generated by one gRNA-mediated editing. More importantly, with one gRNA there remains the possibility of leaving exonic “AG” nucleotides near the cleavage site, which can serve as an alternative pseudo-splice acceptor site.

With its unique T-rich PAM sequence, Cpf1 further expands the genome editing range of the CRISPR family, which is important for potential correction of other disease-related mutations since not all mutation sites contain G-rich PAM sequences for SpCas9 or PAMs for other Cas9 orthologues. Moreover, the staggered cut generated by Cpf1 may be also advantageous for NHEJ-mediated genome editing (Maresca *et al.*, 2013). Finally, the LbCpf1 used in this study is 140-amino-acids smaller than the most widely used SpCas9, which would enhance packaging and delivery by AAV. To evaluate the targeting specificity of Cpf1, two groups (Kim *et al.*, 2016; Tsai *et al.*, 2016) determined the genome-wide editing efficiency of LbCpf1 and AsCpf1 by multiple methods. Both studies showed that LbCpf1 and AsCpf1 had high genome-wide targeting efficiency comparable to that of SpCas9 and high targeting specificity because LbCpf1 and AsCpf1 cannot tolerate mismatches at the 5' PAM proximal region, lessening the frequency of off-targeting effect.

These findings show that Cpf1 is highly efficient in correcting human *DMD* and mouse *Dmd* mutations *in vitro* and *in vivo*. CRISPR-Cpf1-mediated genome editing represents a new and powerful approach to permanently eliminate genetic mutations and rescue abnormalities associated with DMD and other disorders.

* * * * *

All of the compositions and/or methods disclosed and claimed herein can be made and executed without undue experimentation in light of the present disclosure. While the compositions and methods of this disclosure have been described in terms of preferred embodiments, it will be apparent to those of skill in the art that variations may be applied to the compositions and/or methods and in the steps or in the sequence of steps of the method described herein without departing from the concept, spirit and scope of the disclosure. More specifically, it will be apparent that certain agents which are both chemically and physiologically related may be substituted for the agents described herein while the same or similar results would be achieved. All such similar substitutes and modifications apparent to those skilled in the art are deemed to be within the spirit, scope and concept of the disclosure as defined by the appended claims.

VI. References

The following references, to the extent that they provide exemplary procedural or other details supplementary to those set forth herein, are specifically incorporated herein by reference.

Angel *et al.*, *Mol. Cell. Biol.*, 7:2256, 1987a.

Angel *et al.*, *Cell*, 49:729, 1987b.

Aartsma-Rus *et al.*, *Hum. Mutat.* 30, 293–299, 2009.

Baichwal and Sugden, In: *Gene Transfer*, Kucherlapati (Ed), NY, Plenum Press, 117-148, 1986.

Banerji *et al.*, *Cell*, 27(2 Pt 1):299-308, 1981.

Banerji *et al.*, *Cell*, 33(3):729-740, 1983.

Barnes *et al.*, *J. Biol. Chem.*, 272(17):11510-7, 1997.

Baskin *et al.*, *EMBO Mol Med* 6, 1610–1621, 2014.

Benvenisty and Neshif, *Proc. Natl. Acad. Sci. USA*, 83:9551-9555, 1986.

Berkhout *et al.*, *Cell*, 59:273-282, 1989.

Bhavsar *et al.*, *Genomics*, 35(1):11-23, 1996.

Bikard *et al.*, *Nucleic Acids Res.* 41(15): 7429-7437, 2013.

Blanar *et al.*, *EMBO J.*, 8:1139, 1989.

Bodine and Ley, *EMBO J.*, 6:2997, 1987.

Boshart *et al.*, *Cell*, 41:521, 1985.

Bostick *et al.*, *Mol Ther* 19, 1826–1832, 2011.

Bosze *et al.*, *EMBO J.*, 5(7):1615-1623, 1986.

Braddock *et al.*, *Cell*, 58:269, 1989.

Bulla and Siddiqui, *J. Virol.*, 62:1437, 1986.

Burridge *et al.*, *Nat. Methods* 11, 855–860, 2014.

Bushby *et al.*, *Lancet Neurol.*, 9(1): 77-93 (2010).

Bushby *et al.*, *Lancet Neurol.*, 9(2): 177-198 (2010).

Campbell & Kahl, *Nature* 338, 259–262, 1989.

Campbell and Villarreal, *Mol. Cell. Biol.*, 8:1993, 1988.

Campere and Tilghman, *Genes and Dev.*, 3:537, 1989.

Campo *et al.*, *Nature*, 303:77, 1983.

Celander and Haseltine, *J. Virology*, 61:269, 1987.

Celander *et al.*, *J. Virology*, 62:1314, 1988.

Chandler *et al.*, *Cell*, 33:489, 1983.

- Chang *et al.*, *Mol. Cell. Biol.*, 9:2153, 1989.
- Chang *et al.*, *Stem Cells.*, 27:1042-1049, 2009.
- Chatterjee *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:9114, 1989.
- Chen and Okayama, *Mol. Cell Biol.*, 7:2745-2752, 1987.
- Cho *et al.*, *Nat. Biotechnol.* 31(3): 230-232, 2013.
- Choi *et al.*, *Cell*, 53:519, 1988.
- Cirak *et al.*, *The Lancet* 378, 595–605, 2011.
- Coffin, In: *Virology*, Fields *et al.* (Eds.), Raven Press, NY, 1437-1500, 1990.
- Cohen *et al.*, *J. Cell. Physiol.*, 5:75, 1987.
- Costa *et al.*, *Mol. Cell. Biol.*, 8:81, 1988.
- Couch *et al.*, *Am. Rev. Resp. Dis.*, 88:394-403, 1963.
- Coupar *et al.*, *Gene*, 68:1-10, 1988.
- Cripe *et al.*, *EMBO J.*, 6:3745, 1987.
- Culotta and Hamer, *Mol. Cell. Biol.*, 9:1376, 1989.
- Dandolo *et al.*, *J. Virology*, 47:55-64, 1983.
- De Villiers *et al.*, *Nature*, 312(5991):242-246, 1984.
- Deschamps *et al.*, *Science*, 230:1174-1177, 1985.
- DeWitt *et al.*, *Sci Transl Med* 8, 360ra134–360ra134, 2016.
- Donnelly *et al.*, *J. Gen. Virol.* 82, 1027-1041, 2001.
- Dubensky *et al.*, *Proc. Natl. Acad. Sci. USA*, 81:7529-7533, 1984.
- Edbrooke *et al.*, *Mol. Cell. Biol.*, 9:1908, 1989.
- Edlund *et al.*, *Science*, 230:912-916, 1985.
- EP 0273085
- Fechheimer *et al.*, *Proc Natl. Acad. Sci. USA*, 84:8463-8467, 1987.
- Feng and Holland, *Nature*, 334:6178, 1988.
- Ferkol *et al.*, *FASEB J.*, 7:1081-1091, 1993.
- Firak and Subramanian, *Mol. Cell. Biol.*, 6:3667, 1986.
- Foecking and Hofstetter, *Gene*, 45(1):101-105, 1986.
- Fonfara *et al.*, *Nature* 532, 517–521, 2016.
- Fraley *et al.*, *Proc Natl. Acad. Sci. USA*, 76:3348-3352, 1979
- Franz *et al.*, *Cardoscience*, 5(4):235-43, 1994.
- Friedmann, *Science*, 244:1275-1281, 1989.
- Fujita *et al.*, *Cell*, 49:357, 1987.

- Ghosh and Bachhawat, In: *Liver Diseases, Targeted Diagnosis and Therapy Using Specific Receptors and Ligands*, Wu et al. (Eds.), Marcel Dekker, NY, 87-104, 1991.
- Ghosh-Choudhury et al., *EMBO J.*, 6:1733-1739, 1987.
- Gilles et al., *Cell*, 33:717, 1983.
- Gloss et al., *EMBO J.*, 6:3735, 1987.
- Godbout et al., *Mol. Cell. Biol.*, 8:1169, 1988.
- Gomez-Foix et al., *J. Biol. Chem.*, 267:25129-25134, 1992.
- Goncalves et al., *Mol Ther*, 19(7): 1331-1341 (2011).
- Goodbourn and Maniatis, *Proc. Natl. Acad. Sci. USA*, 85:1447, 1988.
- Goodbourn et al., *Cell*, 45:601, 1986.
- Gopal, *Mol. Cell Biol.*, 5:1188-1190, 1985.
- Gopal-Srivastava et al., *J. Mol. Cell. Biol.*, 15(12):7081-90, 1995.
- Graham and Prevec, In: *Methods in Molecular Biology: Gene Transfer and Expression Protocol*, Murray (Ed.), Humana Press, Clifton, NJ, 7:109-128, 1991.
- Graham and van der Eb, *Virology*, 52:456-467, 1973.
- Graham et al., *J. Gen. Virol.*, 36:59-72, 1977.
- Greene et al., *Immunology Today*, 10:272, 1989
- Grosschedl and Baltimore, *Cell*, 41:885, 1985.
- Grunhaus and Horwitz, *Seminar in Virology*, 3:237-252, 1992.
- Harland and Weintraub, *J. Cell Biol.*, 101:1094-1099, 1985.
- Haslinger and Karin, *Proc. Natl. Acad. Sci. USA*, 82:8572, 1985.
- Hauber and Cullen, *J. Virology*, 62:673, 1988.
- Hen et al., *Nature*, 321:249, 1986.
- Hensel et al., *Lymphokine Res.*, 8:347, 1989.
- Hermonat and Muzycska, *Proc. Nat'l Acad. Sci. USA*, 81:6466-6470, 1984.
- Herr and Clarke, *Cell*, 45:461, 1986.
- Hersdorffer et al., *DNA Cell Biol.*, 9:713-723, 1990.
- Herz and Gerard, *Proc. Nat'l. Acad. Sci. USA* 90:2812-2816, 1993.
- Hirochika et al., *J. Virol.*, 61:2599, 1987.
- Holbrook et al., *Virology*, 157:211, 1987.
- Hollinger & Chamberlain, *Current Opinion in Neurology* 28, 522–527, 2015.
- Horlick and Benfield, *Mol. Cell. Biol.*, 9:2396, 1989.
- Horwich et al., *J. Virol.*, 64:642-650, 1990.
- Hsu et al., *Natl Biotechnol.* 31:827-832, 2013

- Huang *et al.*, *Cell*, 27:245, 1981.
- Hug *et al.*, *Mol. Cell. Biol.*, 8:3065, 1988.
- Hwang *et al.*, *Mol. Cell. Biol.*, 10:585, 1990.
- Imagawa *et al.*, *Cell*, 51:251, 1987.
- Imbra and Karin, *Nature*, 323:555, 1986.
- Imler *et al.*, *Mol. Cell. Biol.*, 7:2558, 1987.
- Imperiale and Nevins, *Mol. Cell. Biol.*, 4:875, 1984.
- Jakobovits *et al.*, *Mol. Cell. Biol.*, 8:2555, 1988.
- Jameel and Siddiqui, *Mol. Cell. Biol.*, 6:710, 1986.
- Jaynes *et al.*, *Mol. Cell. Biol.*, 8:62, 1988.
- Jinek *et al.*, *Science* 337, 816-821, 2012.
- Johnson *et al.*, *Mol. Cell. Biol.*, 9:3393, 1989.
- Jones and Shenk, *Cell*, 13:181-188, 1978.
- Kadesch and Berg, *Mol. Cell. Biol.*, 6:2593, 1986.
- Kaneda *et al.*, *Science*, 243:375-378, 1989.
- Karin *et al.*, *Mol. Cell. Biol.*, 7:606, 1987.
- Karlsson *et al.*, *EMBO J.*, 5:2377-2385, 1986.
- Katinka *et al.*, *Cell*, 20:393, 1980.
- Kato *et al.*, *J Biol Chem.*, 266(6):3361-3364, 1991.
- Kawamoto *et al.*, *Mol. Cell. Biol.*, 8:267, 1988.
- Kelly *et al.*, *J. Cell Biol.*, 129(2):383-96, 1995.
- Kiledjian *et al.*, *Mol. Cell. Biol.*, 8:145, 1988.
- Kim *et al.*, *Nature Biotechnology*, 1-2, 2016.
- Kim *et al.*, *Nature Biotechnology* 34, 876-881, 2016.
- Kimura *et al.*, *Dev. Growth Differ.*, 39(3):257-65, 1997.
- Klamut *et al.*, *Mol. Cell. Biol.*, 10:193, 1990.
- Klein *et al.*, *Nature*, 327:70-73, 1987.
- Koch *et al.*, *Mol. Cell. Biol.*, 9:303, 1989.
- Kriegler and Botchan, *In: Eukaryotic Viral Vectors*, Gluzman (Ed.), Cold Spring Harbor: Cold Spring Harbor Laboratory, NY, 1982.
- Kriegler and Botchan, *Mol. Cell. Biol.*, 3:325, 1983.
- Kriegler *et al.*, *Cell*, 38:483, 1984.
- Kriegler *et al.*, *Cell*, 53:45, 1988.
- Kuhl *et al.*, *Cell*, 50:1057, 1987.

- Kunz *et al.*, *Nucl. Acids Res.*, 17:1121, 1989.
- LaPointe *et al.*, *Hypertension*, 27(3):715-22, 1996
- LaPointe *et al.*, *J. Biol. Chem.*, 263(19):9075-8, 1988.
- Larsen *et al.*, *Proc. Natl. Acad. Sci. USA.*, 83:8283, 1986.
- Laspia *et al.*, *Cell*, 59:283, 1989.
- Latimer *et al.*, *Mol. Cell. Biol.*, 10:760, 1990.
- Le Gal La Salle *et al.*, *Science*, 259:988-990, 1993.
- Lee *et al.*, *Nature*, 294:228, 1981.
- Levinson *et al.*, *Nature*, 295:79, 1982.
- Levrero *et al.*, *Gene*, 101:195-202, 1991.
- Lin *et al.*, *Mol. Cell. Biol.*, 10:850, 1990.
- Long *et al.*, *Science* 345:1184-1188, 2014.
- Long *et al.*, Genome Editing of Monogenic Neuromuscular Diseases: A Systematic Review.
JAMA Neurol, doi:10.1001/jamaneurol.2016.3388, 2016.
- Long *et al.*, *Science* 351, 400–403, 2016.
- Luria *et al.*, *EMBO J.*, 6:3307, 1987.
- Lusky and Botchan, *Proc. Natl. Acad. Sci. USA*, 83:3609, 1986.
- Lusky *et al.*, *Mol. Cell. Biol.*, 3:1108, 1983.
- Majors and Varmus, *Proc. Natl. Acad. Sci. USA*, 80:5866, 1983.
- Mali *et al.*, *Science* 339, 823-826, 2013a.
- Mali *et al.*, *Nat Methods* 10, 957-963, 2013b.
- Mali *et al.*, *Nat. Biotechnol.* 31:833-838, 2013c.
- Mann *et al.*, *Cell*, 33:153-159, 1983.
- Maresca *et al.*, *Genome Research* 23, 539–546, 2013.
- Markowitz *et al.*, *J. Virol.*, 62:1120-1124, 1988.
- Martari, M., *et al. Hum Gene Ther*, 20(7):759-766 (2009).
- McNeall *et al.*, *Gene*, 76:81, 1989.
- Miksicek *et al.*, *Cell*, 46:203, 1986.
- Millay *et al.*, *Nat. Med.* 14, 442–447, 2008.
- Mojica *et al.*, *J. Mol. Evol.* 60, 174–182, 2005.
- Mordacq and Linzer, *Genes and Dev.*, 3:760, 1989.
- Moreau *et al.*, *Nucl. Acids Res.*, 9:6047, 1981.
- Moss *et al.*, *J. Gen. Physiol.*, 108(6):473-84, 1996.
- Mourkioti *et al.*, *Nature Cell Biology* 15, 895–904, 2013.

- Muesing *et al.*, *Cell*, 48:691, 1987.
- Nelson *et al.*, *Science* 351, 403–407, 2016.
- Nicolas and Rubinstein, In: *Vectors: A survey of molecular cloning vectors and their uses*, Rodriguez and Denhardt (Eds.), Stoneham: Butterworth, 494-513, 1988.
- Nicolau and Sene, *Biochim. Biophys. Acta*, 721:185-190, 1982.
- Nicolau *et al.*, *Methods Enzymol.*, 149:157-176, 1987.
- Ondek *et al.*, *EMBO J.*, 6:1017, 1987.
- Ornitz *et al.*, *Mol. Cell. Biol.*, 7:3466, 1987.
- Padgett, R.A., *Trends Genet.* 28, 147–154, 2012.
- Palmiter *et al.*, *Cell*, 29:701, 1982a.
- Palmiter *et al.*, *Nature*, 300:611, 1982b.
- Paskind *et al.*, *Virology*, 67:242-248, 1975.
- Pech *et al.*, *Mol. Cell. Biol.*, 9:396, 1989.
- Perales *et al.*, *Proc. Natl. Acad. Sci. USA*, 91(9):4086-4090, 1994.
- Perez-Stable and Constantini, *Mol. Cell. Biol.*, 10:1116, 1990.
- Picard and Schaffner, *Nature*, 307:83, 1984.
- Pinkert *et al.*, *Genes and Dev.*, 1:268, 1987.
- Ponta *et al.*, *Proc. Natl. Acad. Sci. USA*, 82:1020, 1985.
- Porton *et al.*, *Mol. Cell. Biol.*, 10:1076, 1990.
- Potter *et al.*, *Proc. Natl. Acad. Sci. USA*, 81:7161-7165, 1984.
- Queen and Baltimore, *Cell*, 35:741, 1983.
- Quinn *et al.*, *Mol. Cell. Biol.*, 9:4713, 1989.
- Racher *et al.*, *Biotech. Techniques*, 9:169-174, 1995.
- Ragot *et al.*, *Nature*, 361:647-650, 1993.
- Ran *et al.*, *Nature* 520, 186–191, 2015.
- Redondo *et al.*, *Science*, 247:1225, 1990.
- Reisman and Rotter, *Mol. Cell. Biol.*, 9:3571, 1989.
- Renan, *Radiother. Oncol.*, 19:197-218, 1990.
- Resendez Jr. *et al.*, *Mol. Cell. Biol.*, 8:4579, 1988.
- Rich *et al.*, *Hum. Gene Ther.*, 4:461-476, 1993.
- Ridgeway, In: *Vectors: A survey of molecular cloning vectors and their uses*, Stoneham: Butterworth, 467-492, 1988.
- Ripe *et al.*, *Mol. Cell. Biol.*, 9:2224, 1989.
- Rippe *et al.*, *Mol. Cell Biol.*, 10:689-695, 1990.

- Rittling *et al.*, *Nuc. Acids Res.*, 17:1619, 1989.
- Rosen *et al.*, *Cell*, 41:813, 1988.
- Rosenfeld *et al.*, *Cell*, 68:143-155, 1992.
- Rosenfeld *et al.*, *Science*, 252:431-434, 1991.
- Roux *et al.*, *Proc. Natl. Acad. Sci. USA*, 86:9079-9083, 1989.
- Sakai *et al.*, *Genes and Dev.*, 2:1144, 1988.
- Satake *et al.*, *J. Virology*, 62:970, 1988.
- Schaffner *et al.*, *J. Mol. Biol.*, 201:81, 1988.
- Searle *et al.*, *Mol. Cell. Biol.*, 5:1480, 1985.
- Sharp and Marciniak, *Cell*, 59:229, 1989.
- Shaul and Ben-Levy, *EMBO J.*, 6:1913, 1987.
- Sherman *et al.*, *Mol. Cell. Biol.*, 9:50, 1989.
- Shimizu-Motohashi *et al.*, *Am J Transl Res* 8, 2471–2489, 2016.
- Sleigh and Lockett, *J. EMBO*, 4:3831, 1985.
- Spalholz *et al.*, *Cell*, 42:183, 1985.
- Spandau and Lee, *J. Virology*, 62:427, 1988.
- Spandidos and Wilkie, *EMBO J.*, 2:1193, 1983.
- Stephens and Hentschel, *Biochem. J.*, 248:1, 1987.
- Stratford-Perricaudet and Perricaudet, In: *Human Gene Transfer*, Cohen-Haguenauer and Boiron (Eds.), John Libbey Eurotext, France, 51-61, 1991.
- Stratford-Perricaudet *et al.*, *Hum. Gene. Ther.*, 1:241-256, 1990.
- Stuart *et al.*, *Nature*, 317:828, 1985.
- Sullivan and Peterlin, *Mol. Cell. Biol.*, 7:3315, 1987.
- Swartzendruber and Lehman, *J. Cell. Physiology*, 85:179, 1975.
- Tabebordbar *et al.*, *Science* 351, 407–411, 2016.
- Takebe *et al.*, *Mol. Cell. Biol.*, 8:466, 1988.
- Tavernier *et al.*, *Nature*, 301:634, 1983.
- Taylor and Kingston, *Mol. Cell. Biol.*, 10:165, 1990a.
- Taylor and Kingston, *Mol. Cell. Biol.*, 10:176, 1990b.
- Taylor *et al.*, *J. Biol. Chem.*, 264:15160, 1989.
- Temin, In: *Gene Transfer*, Kucherlapati (Ed.), NY, Plenum Press, 149-188, 1986.
- Thiesen *et al.*, *J. Virology*, 62:614, 1988.
- Top *et al.*, *J. Infect. Dis.*, 124:155-160, 1971.
- Tóth *et al.*, *Biology Direct*, 1–14, 2016.

- Tronche *et al.*, *Mol. Biol. Med.*, 7:173, 1990.
- Trudel and Constantini, *Genes and Dev.* 6:954, 1987.
- Tsai *et al.*, *Nature Biotechnology* 34, 882–887, 2016.
- Tur-Kaspa *et al.*, *Mol. Cell Biol.*, 6:716-718, 1986.
- Tyndell *et al.*, *Nuc. Acids. Res.*, 9:6231, 1981.
- Vannice and Levinson, *J. Virology*, 62:1305, 1988.
- Varmus *et al.*, *Cell*, 25:23-36, 1981.
- Vasseur *et al.*, *Proc Natl. Acad. Sci. U.S.A.*, 77:1068, 1980.
- Wagner *et al.*, *Proc. Natl. Acad. Sci. USA*, 87(9):3410-3414, 1990.
- Wang *et al.*, *Cell*, 153:910-910, 2013.
- Weber *et al.*, *Cell*, 36:983, 1984.
- Weinberger *et al.* *Mol. Cell. Biol.*, 8:988, 1984.
- Winoto and Baltimore, *Cell*, 59:649, 1989.
- Wong *et al.*, *Gene*, 10:87-94, 1980.
- Wu and Wu, *Adv. Drug Delivery Rev.*, 12:159-167, 1993.
- Wu and Wu, *Biochemistry*, 27:887-892, 1988.
- Wu and Wu, *J. Biol. Chem.*, 262:4429-4432, 1987.
- Wu *et al.*, *Cell Stem Cell* 13, 659-662, 2013.
- Wu *et al.*, *Nat Biotechnol* 32, 670-676, 2014.
- Xu *et al.*, *Mol Ther* 24, 564–569, 2016.
- Yamauchi-Takahara *et al.*, *Proc. Natl. Acad. Sci. USA*, 86(10):3504-3508, 1989.
- Yang *et al.*, *Proc. Natl. Acad. Sci. USA*, 87:9568-9572, 1990.
- Yin *et al.*, *Nat Biotechnol* 32, 551-553, 2014.
- Yin *et al.*, *Physiol Rev* 93, 23-67, 2013.
- Young *et al.*, *Cell Stem Cell* 18, 533–540, 2016.
- Yutzey *et al.* *Mol. Cell. Biol.*, 9:1397, 1989.
- Zechner *et al.*, *Cell Metabolism* 12, 633–642, 2010.
- Zelenin *et al.*, *FEBS Lett.*, 280:94-96, 1991.
- Zetsche *et al.*, *Cell* 163, 759–771, 2015.
- Ziober and Kramer, *J. Bio. Chem.*, 271(37):22915-22922, 1996.

CLAIMS

1. A composition comprising a sequence encoding a Cpf1 polypeptide and a sequence encoding a DMD guide RNA (gRNA), wherein the DMD gRNA targets a dystrophin splice site, and wherein the DMD gRNA comprises any one of SEQ ID No. 448 to 770.
2. The composition of claim 1, wherein the sequence encoding the Cpf1 polypeptide is isolated or derived from a sequence encoding a *Lachnospiraceae* Cpf1 polypeptide .
3. The composition of claim 1, wherein the sequence encoding the Cpf1 polypeptide is isolated or derived from a sequence encoding an *Acidaminococcus* Cpf1 polypeptide.
4. The composition of any one of claims 1-3, wherein the sequence encoding the Cpf1 polypeptide or the sequence encoding the DMD gRNA comprises an RNA sequence.
5. The composition of claim 4, wherein the RNA sequence is an mRNA sequence.
6. The composition of claim 4 or 5, wherein the RNA sequence comprises at least one chemically-modified nucleotide.
7. The composition of any one of claims 1-3, wherein the sequence encoding the Cpf1 polypeptide comprises a DNA sequence.
8. The composition of any one of claims 1-7, wherein a first vector comprises the sequence encoding the Cpf1 polypeptide and a second vector comprises the sequence encoding the DMD gRNA.
9. The composition of claim 8, wherein the first vector or the sequence encoding the Cpf1 polypeptide further comprises a first polyA sequence.
10. The composition of claim 8, wherein the second vector or the sequence encoding the DMD gRNA further comprises a second polyA sequence.

11. The composition of any one of claims 8-10, wherein the first vector or the sequence encoding the Cpf1 polypeptide further comprises a first promoter sequence.
12. The composition of any one of claims 8-10, wherein the second vector or the sequence encoding the DMD gRNA further comprises a second promoter sequence.
13. The composition of claim 11 or 12, wherein the first promoter sequence and the second promoter sequence are identical.
14. The composition of claim 11 or 12, wherein the first promoter sequence and the second promoter sequence are not identical.
15. The composition of any one of claims 11-14, wherein the first promoter sequence or the second promoter sequence comprises a constitutive promoter.
16. The composition of any one of claims 11-14, wherein the first promoter sequence or the second promoter sequence comprises an inducible promoter.
17. The composition of any one of claims 11-16, wherein the first promoter sequence or the second promoter sequence comprises a muscle-cell specific promoter.
18. The composition of claim 17, wherein the muscle-cell specific promoter is a myosin light chain-2 promoter, an α -actin promoter, a troponin 1 promoter, a $\text{Na}^+/\text{Ca}^{2+}$ exchanger promoter, a dystrophin promoter, an $\alpha 7$ integrin promoter, a brain natriuretic peptide promoter, an α B-crystallin/small heat shock protein promoter, an α -myosin heavy chain promoter, or an ANF promoter.
19. The composition of claim 17, wherein the muscle-cell specific promoter is a dystrophin promoter.
20. The composition of claim 17, wherein the muscle-cell specific promoter is an α -myosin heavy chain promoter.

21. The composition of any one of claims 8-20, wherein the first vector or the second vector further comprises a sequence encoding a detectable marker.
22. The composition of claim 21, wherein the detectable marker is a fluorescent marker.
23. The composition of any one of claims 8-22, wherein the first vector or the second vector further comprises a sequence encoding 2A-like self-cleaving domain.
24. The composition of claim 23, wherein the sequence encoding 2A-like self-cleaving domain comprises a TaV-2A peptide.
25. The composition of any one of claims 1-7, wherein vector comprises the sequence encoding the Cpf1 polypeptide and the sequence encoding the DMD gRNA.
26. The composition of claim 25, wherein the vector further comprises a poly A sequence.
27. The composition of claim 25 or 26, wherein the vector further comprises a promoter sequence.
28. The composition of claim 27, wherein the promoter sequence comprises a constitutive promoter.
29. The composition of claim 27, wherein the promoter sequence comprises an inducible promoter.
30. The composition of claim 28 or 29, wherein the promoter sequence comprises a muscle-cell specific promoter.
31. The composition of claim 30, wherein the muscle-cell specific promoter is a myosin light chain-2 promoter, an α -actin promoter, a troponin 1 promoter, a $\text{Na}^+/\text{Ca}^{2+}$ exchanger promoter, a dystrophin promoter, an $\alpha 7$ integrin promoter, a brain natriuretic peptide promoter, an α B-crystallin/small heat shock protein promoter, an α -myosin heavy chain promoter, or an ANF promoter.

32. The composition of claim 30, wherein the muscle-cell specific promoter is a dystrophin promoter.
33. The composition of claim 30, wherein the muscle-cell specific promoter is an α -myosin heavy chain promoter.
34. The composition of any one of claims 25-33, wherein the first vector or the second vector further comprises a sequence encoding a detectable marker.
35. The composition of claim 34, wherein the detectable marker is a fluorescent maker.
36. The composition of any one of claims 25-35, wherein the first vector or the second vector further comprises a sequence encoding 2A-like self-cleaving domain.
37. The composition of claim 36, wherein the sequence encoding 2A-like self-cleaving domain comprises a TaV-2A peptide.
38. The composition of any one of claims 1-37, wherein the composition comprises a sequence codon optimized for expression in a mammalian cell.
39. The composition of any one of claims 1-38, wherein the composition comprises a sequence codon optimized for expression in a human cell.
40. The composition of claim 39, wherein the sequence encoding the Cpf1 polypeptide is codon optimized for expression in human cells.
41. The composition of any one of claims 1-40, wherein the splice site is a splice donor site.
42. The composition of any one of claims 1-40, wherein the splice site is a splice acceptor site.
43. The composition of any one of claims 8-24, wherein the first vector or the second vector is a non-viral vector.

44. The composition of claim 43, wherein the non-viral vector is a plasmid.
45. The composition of claim 43 or 44, wherein a liposome or a nanoparticle comprises the first vector or the second vector.
46. The composition of any one of claims 25-37, wherein the vector is a non-viral vector.
47. The composition of claim 46, wherein the non-viral vector is a plasmid.
48. The composition of claim 46 or 47, wherein a liposome or a nanoparticle comprises the vector.
49. The composition of any one of claims 8-24, wherein the first vector or the second vector is a viral vector.
50. The composition of any one of claims 25-37, wherein the vector is a viral vector.
51. The composition of claim 49 or 50, wherein the viral vector is an adeno-associated viral (AAV) vector.
52. The composition of claim 51, wherein the AAV vector is replication-defector or conditionally replication defective.
53. The composition of claim 51 or 52, wherein the AAV vector is a recombinant AAV vector.
54. The composition of any one of claims 51-53, wherein the AAV vector comprises a sequence isolated or derived from an AAV vector of serotype AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11 or any combination thereof.
55. The composition of any one of claims 1-54, further comprising a single-stranded DMD oligonucleotide.

56. The composition of any one of claims 1-55, further comprising a pharmaceutically acceptable carrier.
57. A cell comprising the composition of any one of claims 1-56.
58. The cell of claim 57, wherein the cell is a muscle cell, a satellite cell or a precursor thereof.
59. The cell of claim 57, wherein the cell is an iPSC or an iCM.
60. A composition comprising the cell of any one of claims 57-59.
61. A method of correcting a dystrophin gene defect comprising contacting a cell and a composition of any one of claims 1-59 under conditions suitable for expression of the Cpf1 polypeptide and the gRNA, wherein the Cpf1 polypeptide disrupts the dystrophin splice site; and wherein disruption of the splice site results in selective skipping of a mutant DMD exon.
62. The method of claim 61, wherein the mutant DMD exon is exon 23.
63. The method of claim 61, wherein the mutant DMD exon is exon 51.
64. The method of any one of claims 61-63, wherein the cell is in vivo, ex vivo, in vitro or in situ.
65. A cell produced by the method of any one of claims 61-64.
66. A composition comprising the cell of claim 65.
67. A method of treating muscular dystrophy in a subject in need thereof comprising administering to the subject a therapeutically effective amount of a composition of any one of claims 1-56, 60, or 66.
68. The method of claim 67, wherein the composition is administered locally.

69. The method of claim 67 or 68, wherein the composition is administered directly to a muscle tissue.

70. The method of claim 69, wherein the composition is administered by intramuscular infusion or injection.

71. The method of claim 69 or 70, wherein the muscle tissue comprises a tibialis anterior tissue, a quadricep tissue, a soleus tissue, a diaphragm tissue or a heart tissue.

72. The method of claim 67, wherein the composition is administered systemically, such as by intravenous infusion or injection.

73. The method of any one of claims 67-72, wherein, following administration of the composition, the subject exhibits normal dystrophin-positive myofibers, and mosaic dystrophin-positive myofibers containing centralized nuclei, or a combination thereof.

74. The method of any one of claims 67-72, wherein, following administration of the composition, the subject exhibits an emergence or an increase in a level of abundance of normal dystrophin-positive myofibers when compared to an absence or an level of abundance of normal dystrophin-positive myofibers prior to administration of the composition.

75. The method of any one of claims 67-72, wherein, following administration of the composition, the subject exhibits an emergence or an increase in a level of abundance of mosaic dystrophin-positive myofibers containing centralized nuclei when compared to an absence or an level of abundance of mosaic dystrophin-positive myofibers containing centralized nuclei prior to administration of the composition.

76. The method of any one of claims 67-72, wherein, following administration of the composition, the subject exhibits a decreased serum CK level when compared to a serum CK level prior to administration of the composition.

77. The method of any one of claims 67-72, wherein, following administration of the composition, the subject exhibits improved grip strength when compared to a grip strength prior to administration of the composition.

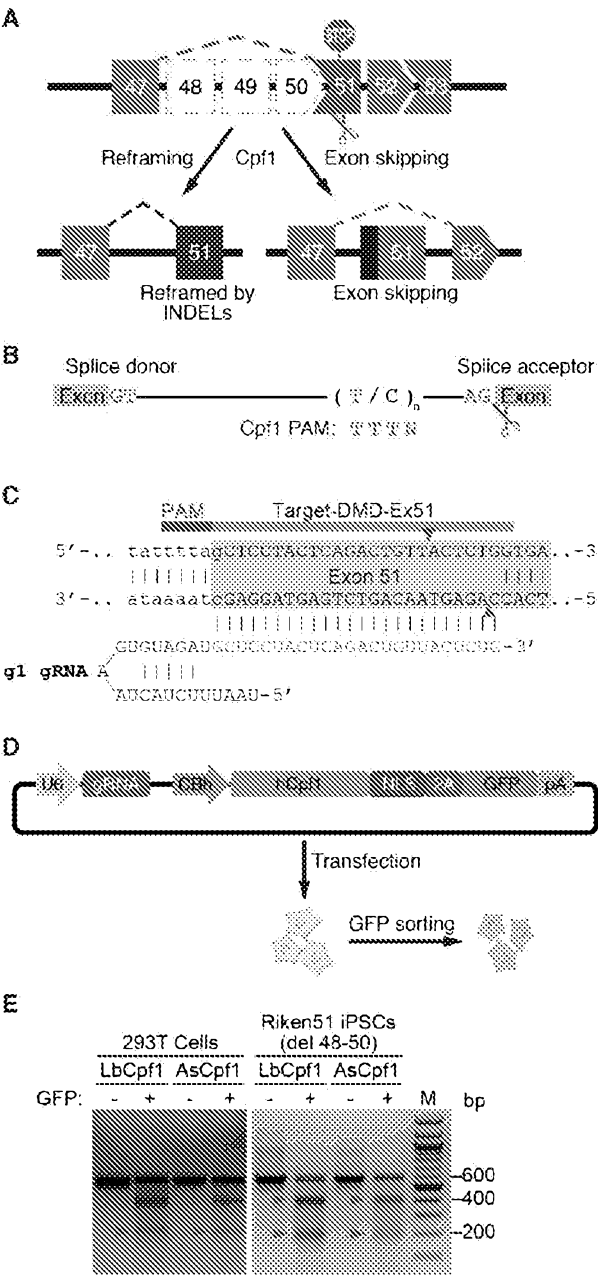
78. The method of any one of claims 67-77, wherein the method comprises administering a therapeutically effective amount of a composition of claim 60 or 66, wherein the cell is autologous.
79. The method of any one of claims 67-77, wherein the method comprises administering a therapeutically effective amount of a composition of claim 60 or 66, wherein the cell is allogeneic.
80. The method of any one of claims 67-79, wherein the subject is a neonate, an infant, a child, a young adult, or an adult.
81. The method of any one of claims 67-80, wherein the subject has muscular dystrophy.
82. The method of any one of claims 67-80, wherein the subject is a genetic carrier for muscular dystrophy.
83. The method of any one of claims 67-82, wherein the subject is male.
84. The method of any one of claims 67-82, wherein the subject is female.
85. The method of any one of claims 67-84, wherein the subject appears to be asymptomatic and wherein a genetic diagnosis reveals a mutation in one or both copies of a *DMD* gene that impairs function of the *DMD* gene product.
86. The method of any one of claims 67-84, wherein the subject presents an early sign or symptom of muscular dystrophy.
87. The method of claim 86, wherein the early sign or symptom of muscular dystrophy comprises loss of muscle mass or proximal muscle weakness.
88. The method of claim 87, wherein the loss of muscle mass or proximal muscle weakness occurs in one or both leg(s) and/or a pelvis, followed by one or more upper body muscle(s).

89. The method of claim 86, wherein the early sign or symptom of muscular dystrophy further comprises pseudohypertrophy, low endurance, difficulty standing, difficulty walking, difficulty ascending a staircase or a combination thereof.
90. The method of any one of claims 67-84, wherein the subject presents a progressive sign or symptom of muscular dystrophy.
91. The method of claim 90, wherein the progressive sign or symptom of muscular dystrophy comprises muscle tissue wasting, replacement of muscle tissue with fat, or replacement of muscle tissue with fibrotic tissue.
92. The method of any one of claims 67-84, wherein the subject presents a later sign or symptom of muscular dystrophy.
93. The method of claim 92, wherein the later sign or symptom of muscular dystrophy comprises abnormal bone development, curvature of the spine, loss of movement, and paralysis.
94. The method of any one of claims 67-84, wherein the subject presents a neurological sign or symptom of muscular dystrophy.
95. The method of claim 94, wherein the neurological sign or symptom of muscular dystrophy comprises intellectual impairment and paralysis.
96. The method of any one of claims 67-95, wherein the administration of the composition occurs prior to the subject presenting one or more progressive, later or neurological signs or symptoms of muscular dystrophy.
97. The method of any one of claims 67-96, wherein the subject is less than 10 years old.
98. The method of claim 97, wherein the subject is less than 5 years old.
99. The method of claim 97, wherein the subject is less than 2 years old.

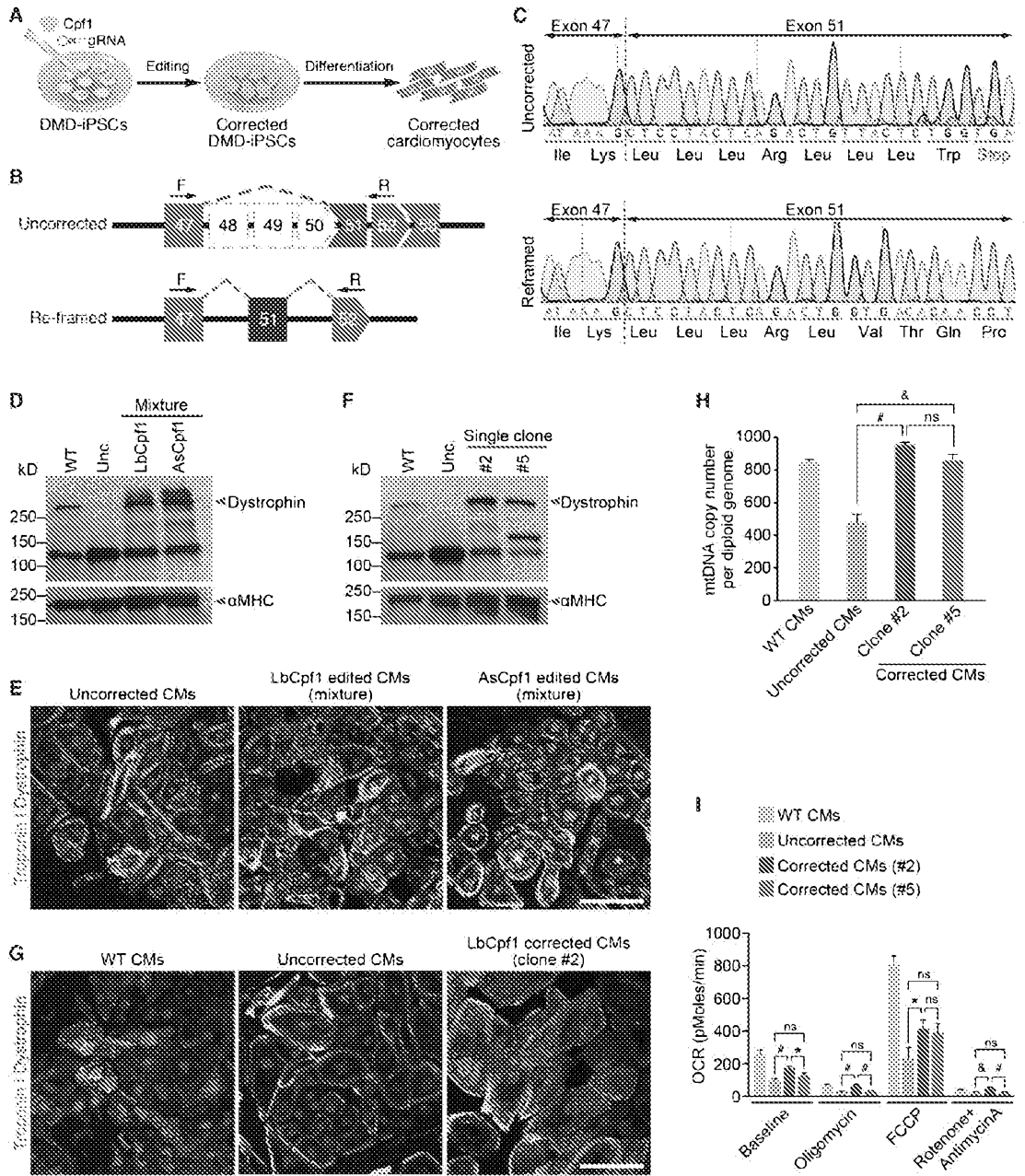
100. Use of a therapeutically-effective amount of the composition of any one of claims 1-56, 60, or 66 for treating muscular dystrophy in a subject in need thereof.

101. Use of a therapeutically-effective amount of the composition of claim 60 or 66 for treating muscular dystrophy in a subject in need thereof, wherein the cell is autologous.

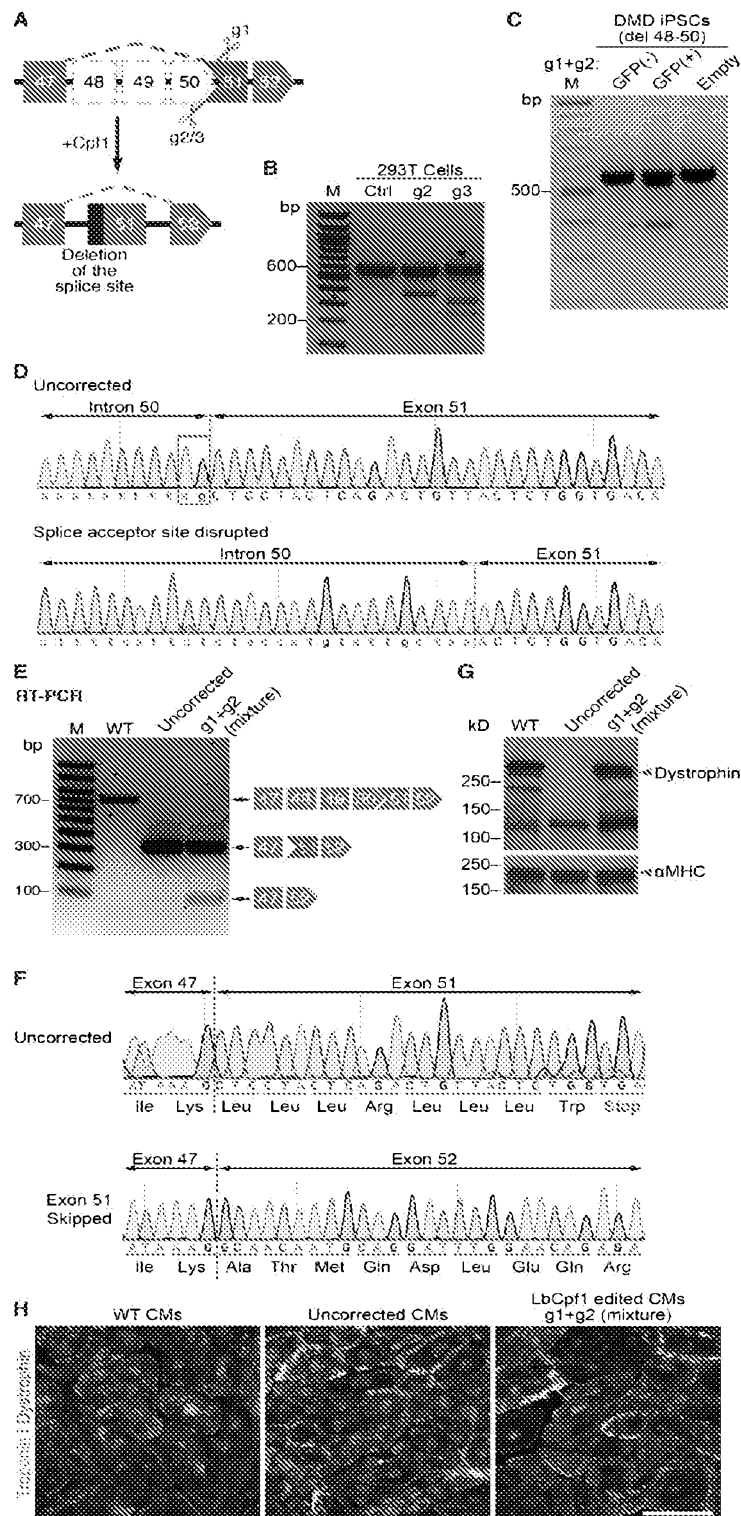
102. Use of a therapeutically-effective amount of the composition of claim 60 or 66 for treating muscular dystrophy in a subject in need thereof, wherein the cell is allogeneic.



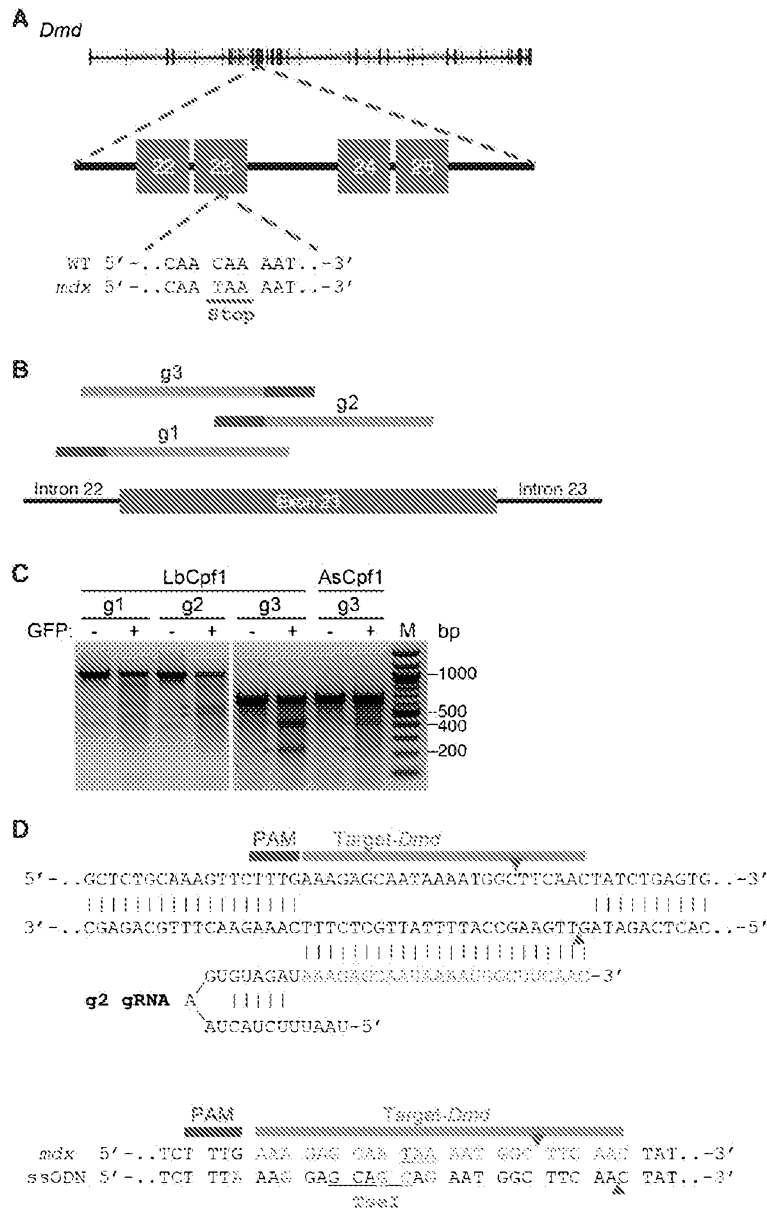
FIGS. 1A-E



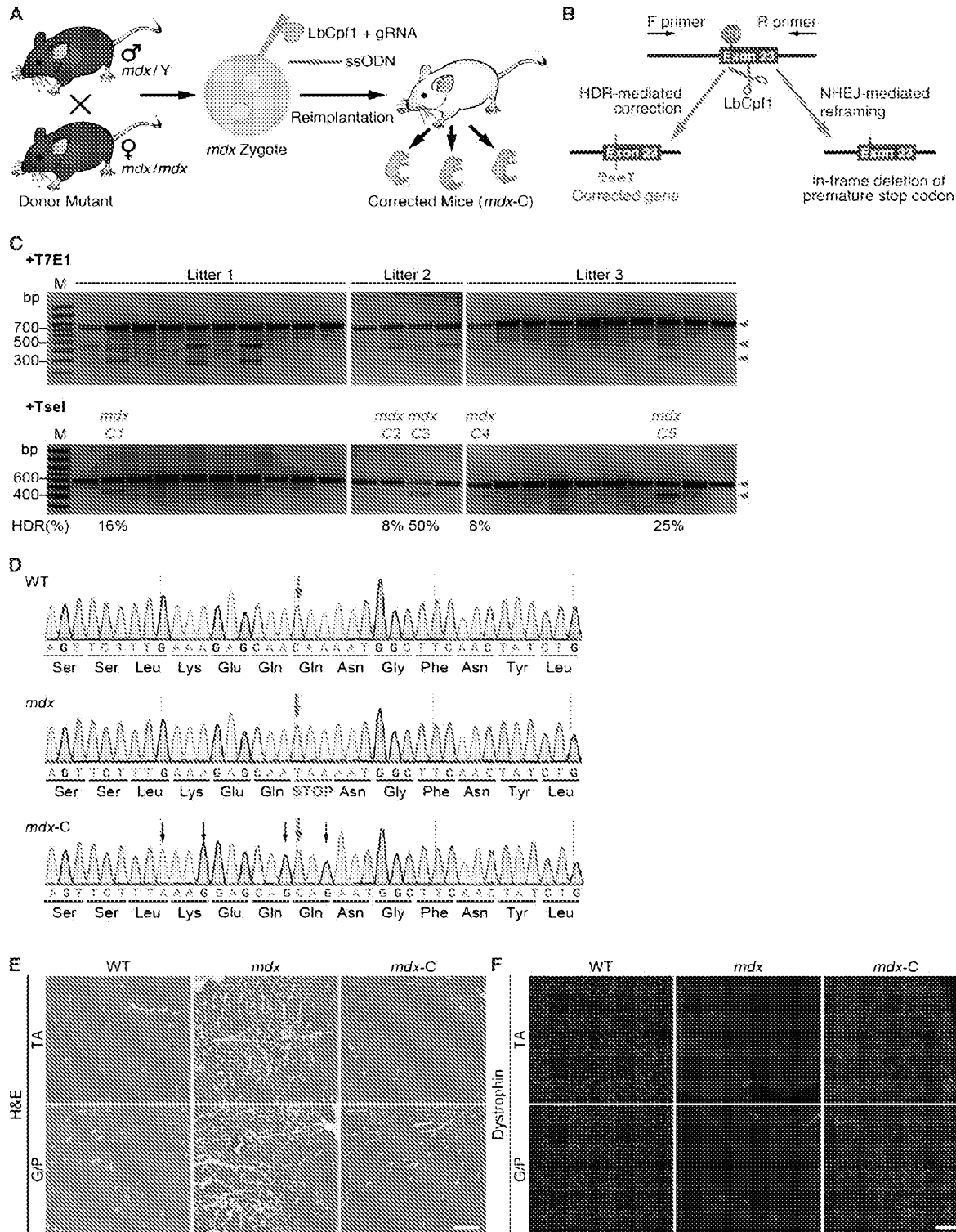
FIGS. 2A-I



FIGS. 3A-H



FIGS. 4A-D



FIGS. 5A-F

A**Riken 51 iPSCs mixtures edited by LbCpf1**

uncor. tatatttagCTCCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTT (SEQ ID NO: 60)
 #01 tatatttagCTCCTACTCAGAC---TACTCTGGTGACACAACCTGTGGTT (Δ3) (SEQ ID NO: 61)
 #02 tatatttagCT-----ACAACCTGTGGTT (Δ26) (In-frame) (SEQ ID NO: 62)
 #03 tatatttagCTCCTACTCAGTA-----GGTGACANAACCTGTGGTT (Δ9) (SEQ ID NO: 63)
 #04 tatatttagCTCCTACTCAGAC---TACTCTGGTGACACAACCTGTGGTT (Δ3) (SEQ ID NO: 64)
 #05 tatatttagCTCCTACTCA-----GACACAACCTGTGGTT (Δ15) (SEQ ID NO: 65)
 #06 tatatttagCTCCTACTC-----ACTCTGGTGACACAACCTGTGGTT (Δ8) (In-frame) (SEQ ID NO: 66)
 #07 tatatttagCTCCTACTCAGAC---TACTCTGGTGACACAACCTGTGGTT (Δ3) (SEQ ID NO: 67)
 #08 tatatttagCTCCTACTCAGACT-----GTGACACAACCTGTGGTT (Δ9) (SEQ ID NO: 68)
 #09 tatatttagCTCCTACTCAGACTG-----TGGTGACACAACCTGTGGTT (Δ6) (SEQ ID NO: 69)
 #10 tatatttagTTCC-----ACTCTGGTGACACAACCTGTGGTT (Δ13) (SEQ ID NO: 70)

Riken 51 iPSCs mixtures edited by AsCpf1

uncor. tatatttagCTCCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTT (SEQ ID NO: 60)
 #01 tatatttagCTCCTACTCAG-----ACTCTGGTGACACAACCTGTGGTT (Δ6) (SEQ ID NO: 71)
 #02 tatatttagCTCCTACTCA-----GTGACACAACCTGTGGTT (Δ13) (SEQ ID NO: 72)
 #03 tatatttagCTCCTACTCAG-----ACTCTGGTGACACAACCTGTGGTT (Δ6) (SEQ ID NO: 73)
 #04 uncor.
 #05 tatatttagCTCCTACTCAGACTGT-----GTGACACAACCTGTGGTT (Δ7) (SEQ ID NO: 74)
 #06 tatatttagCTCCTACTCGG-----GACACAACCTGTGGTT (Δ14) (In-frame) (SEQ ID NO: 75)
 #07 tatatttagCTCCCACTCAGGTAGGACACAACCTGTGGTTACTAAGTAAC (SEQ ID NO: 76)
 #08 tatatttagCTCCTACTCAGA-----CTGGTGACACAACCTGTGGTT (Δ8) (In-frame) (SEQ ID NO: 77)
 #09 tatatttagCTCCTACTCA-----TACTAAGGAACTGTGGTT (Δ12) (SEQ ID NO: 78)
 #10 tatatttagCTCCTACTCAG-----ACTCTGGTGACACAACCTGTGGTT (Δ6) (SEQ ID NO: 79)

B**Riken 51 iPSCs single clones edited by LbCpf1**

uncor.
 #01 tatatttagCTCCTACTCAGACT-----TGGTGACACAACCTGTGGTT (Δ7) (SEQ ID NO: 80)
 #02 tatatttagCTCCTACTCAGA-----CTGGTGACACAACCTGTGGTT (Δ8) (In-frame) (SEQ ID NO: 81)
 #03 tatatttagCTCCTACTCAGA-----CTGTGGTT (Δ21) (SEQ ID NO: 82)
 #04 tatatttagCTCCTACTCAGAC---TACTCTGGTGACACAACCTGTGGTT (Δ3) (SEQ ID NO: 83)
 #05 tatatttagCTC-----CTGGTGACACAACCTGTGGTT (Δ17) (In-frame) (SEQ ID NO: 84)
 #06 tatatttagCTCCTACTCAGAC-----ACACAACCTGTGGTT (Δ13) (SEQ ID NO: 85)
 #07 tatatttagCTCCTACTCAG-----ACTCTGGTGACACAACCTGTGGTT (Δ6) (SEQ ID NO: 86)

Riken 51 iPSCs single clones edited by AsCpf1

uncor. tatatttagCTCCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTT (SEQ ID NO: 60)
 #01 uncor.
 #02 tatatttagCTCCTACTCAG-----ACTCTGGTGACACAACCTGTGGTT (Δ6) (SEQ ID NO: 87)
 #03 tatatttagCTCCTACTCAGA-----CTGGTGACACAACCTGTGGTT (Δ8) (In-frame) (SEQ ID NO: 88)
 #04 tatatttagCTCCTACTCAG-----TGGTGACACAACCTGTGGTT (Δ8) (In-frame) (SEQ ID NO: 89)
 tatatttagCTCCTACTCAGACTGTTACTCTGGTGACACAACCTGTGGTT (SEQ ID NO: 60)

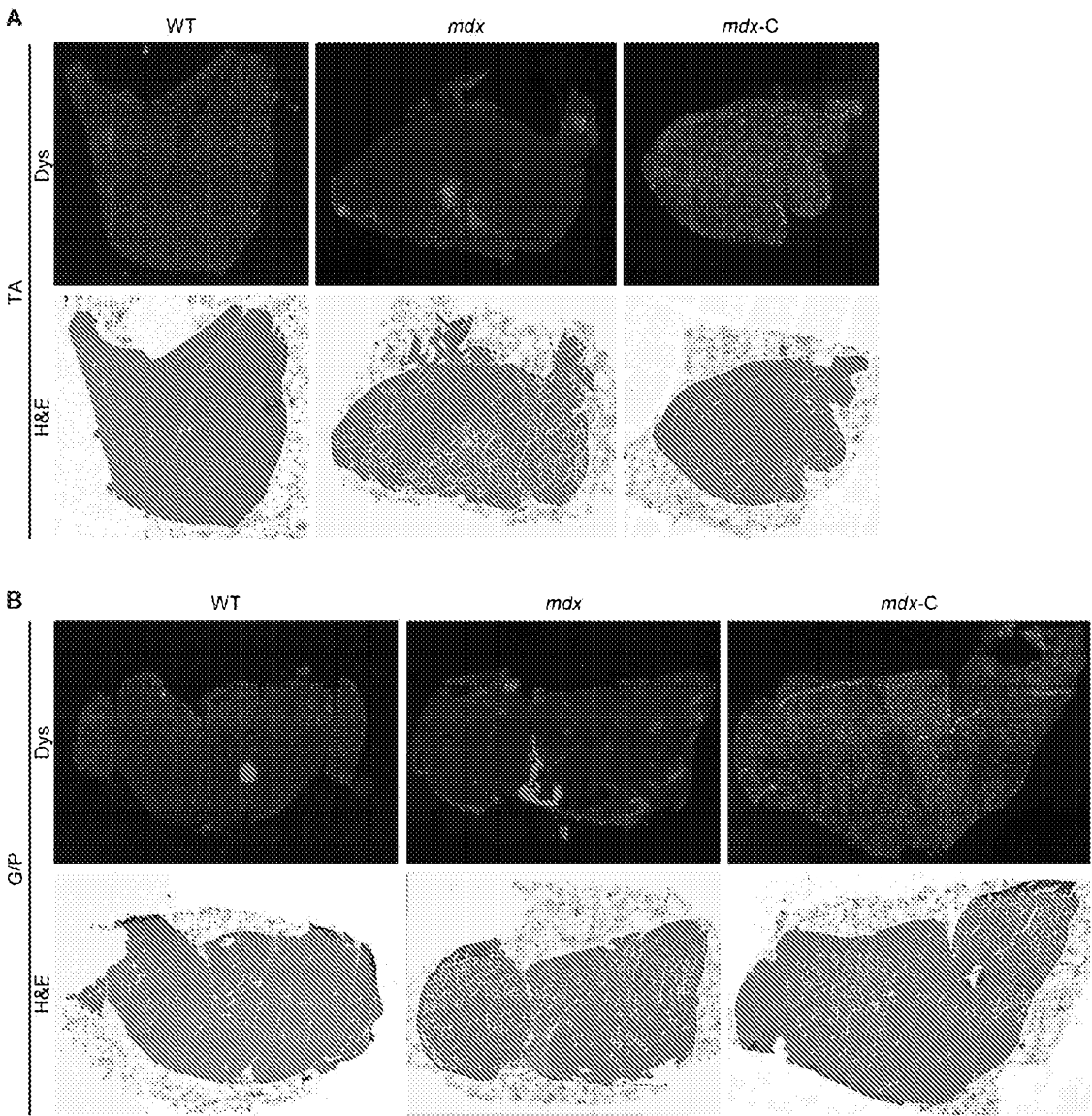
C**Mouse 10T1/2 fibroblasts edited by LbCpf1 with g2 and g3****Dmd-E23-g2**

WT GCAAAGTTCTTTGAAAGAGCAACAAAATGGCTTCAACTATCTGAGTGACACT (SEQ ID NO: 90)
 GCAAAGTTCTTTGAAAGAGCAACAAAATGG--TCAACTATCTGAGTGACACT (SEQ ID NO: 91)
 GCAAAGTTCTTTGAAAGAGCAACAAAATGG-----AACTATCTGAGTGACACT (SEQ ID NO: 92)
 GCAAAGTTCTTTGAAAGAGCAACAAAATGG-----A---TATCTGAGTGACACT (SEQ ID NO: 93)
 GCAAAGTTCTTTGAAAGAG-----CTATCTGAGTGACACT (SEQ ID NO: 94)

Dmd-E23-g3

WT GTGTAAATGAAGTTCTATTTAATTTTGAGGCTCTGCAAAGTTCTTTGAAAGA (SEQ ID NO: 95)
 GTGTAAATGAAGTTCTATT-----G-AAAGTTCTTTGAAAGA (SEQ ID NO: 96)

FIGS. 6A-C



FIGS. 7A-B

SEQUENCE LISTING

<110> THE BOARD OF REGENTS OF THE UNIVERSITY OF TEXAS SYSTEM

<120> PREVENTION OF MUSCULAR DYSTROPHY BY CRISPR/CPFL-MEDIATED GENE EDITING

<130> 332150-2001

<140> Not Yet Assigned

<141> 2017-11-28

<150> 62/426,853

<151> 2016-11-28

<160> 770

<170> PatentIn version 3.5

<210> 1

<400> 1
000

<210> 2

<400> 2
000

<210> 3

<400> 3
000

<210> 4

<400> 4
000

<210> 5

<400> 5
000

<210> 6

<400> 6
000

<210> 7

<400> 7
000

<210> 8

<400> 8
000

<210> 9

<400> 9
000

<210> 10

<400> 10
000

<210> 11

<400> 11
000

<210> 12

<400> 12
000

<210> 13

<400> 13
000

<210> 14

<400> 14
000

<210> 15

<400> 15
000

<210> 16

<400> 16
000

<210> 17

<400> 17
000

<210> 18

<400> 18
000

<210> 19

<400> 19
000

<210> 20

<400> 20
000

<210> 21

<400> 21
000

<210> 22

<400> 22
000

<210> 23

<400> 23
000

<210> 24

<400> 24
000

<210> 25

<400> 25
000

<210> 26

<400> 26
000

<210> 27

<400> 27
000

<210> 28

<400> 28
000

<210> 29

<400> 29
000

<210> 30

<400> 30
000

<210> 31

<400> 31
000

<210> 32

<400> 32
000

<210> 33

<400> 33
000

<210> 34

<400> 34
000

<210> 35

<400> 35
000

<210> 36

<400> 36
000

<210> 37

<400> 37
000

<210> 38

<400> 38
000

<210> 39

<400> 39
000

<210> 40

<400> 40
000

<210> 41

<400> 41
000

<210> 42

<400> 42
000

<210> 43

<400> 43
000

<210> 44

<400> 44
000

<210> 45

<400> 45
000

<210> 46

<400> 46
000

<210> 47

<400> 47
000

<210> 48

<400> 48
000

<210> 49

<400> 49
000

<210> 50

<400> 50
000

<210> 51

<400> 51
000

<210> 52

<400> 52
000

<210> 53

<400> 53
000

<210> 54

<400> 54
000

<210> 55

<400> 55
000

<210> 56

<400> 56
000

<210> 57

<400> 57
000

<210> 58

<400> 58
000

<210> 59

<400> 59
000

<210> 60

<211> 24

<212> DNA

<213> Homo sapiens

<400> 60

tcttttttctt ctttttttctt tttt

24

<210> 61

<211> 24
 <212> DNA
 <213> Homo sapiens

<400> 61
 cttttttcttc ttttttcctt ttg 24

<210> 62
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 62
 tttttcttct tttttccttt ttgc 24

<210> 63
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 63
 tctttctttt tcctttttgc aaaa 24

<210> 64
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 64
 cttctttttt cctttttgca aaaa 24

<210> 65
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 65
 ttcttttttc ctttttgcaa aaac 24

<210> 66
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 66
 ttcttttttg caaaaaccca aaat 24

<210> 67
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 67
 tcctttttgc aaaaacccaa aata 24

<210> 68
 <211> 24

<212> DNA
 <213> Homo sapiens

<400> 68
 ccttttttgca aaaacccaaa atat 24

<210> 69
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 69
 cttttttgcaa aaacccaaaa tatt 24

<210> 70
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 70
 tgcaaaaacc caaaatatttt tagc 24

<210> 71
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 71
 gcaaaaaccc aaaatatttt agct 24

<210> 72
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 72
 caaaaacca aaatattttta gctc 24

<210> 73
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 73
 agtcctact cagactgtta ctct 24

<210> 74
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 74
 gctcctactc agactgttac tctg 24

<210> 75
 <211> 24
 <212> DNA

<213> Homo sapiens	
<400> 75	
cttagtaacc acaggttggtg tcac	24
<210> 76	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 76	
gagatggcag tttccttagt aacc	24
<210> 77	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 77	
tagtttggag atggcagttt cctt	24
<210> 78	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 78	
ttctcataacc ttctgcttga tgat	24
<210> 79	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 79	
tcatttttttc tcataccttc tgct	24
<210> 80	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 80	
atcattttttt ctcatacctt ctgc	24
<210> 81	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 81	
aagaaaaact tctgccaact tttta	24
<210> 82	
<211> 24	
<212> DNA	
<213> Homo sapiens	

<400> 82
aaagaaaaac ttctgccaac tttt 24

<210> 83
<211> 24
<212> DNA
<213> Homo sapiens

<400> 83
tctttaaaat gaagattttc cacc 24

<210> 84
<211> 24
<212> DNA
<213> Homo sapiens

<400> 84
ctttaaaatg aagattttcc acca 24

<210> 85
<211> 24
<212> DNA
<213> Homo sapiens

<400> 85
tttaaaatga agattttcca ccaa 24

<210> 86
<211> 24
<212> DNA
<213> Homo sapiens

<400> 86
aatgaagat tttccaccaa tcac 24

<210> 87
<211> 24
<212> DNA
<213> Homo sapiens

<400> 87
ccaccaatca ctttactctc ctag 24

<210> 88
<211> 24
<212> DNA
<213> Homo sapiens

<400> 88
caccaatcac tttactctcc taga 24

<210> 89
<211> 24
<212> DNA
<213> Homo sapiens

<400> 89 ctctcctaga ccatttccca ccag	24
<210> 90 <211> 24 <212> DNA <213> Homo sapiens	
<400> 90 agaaaagatt aaacagtgtg ctac	24
<210> 91 <211> 24 <212> DNA <213> Homo sapiens	
<400> 91 tttgagaaaa gattaaacag tgtg	24
<210> 92 <211> 24 <212> DNA <213> Homo sapiens	
<400> 92 atttgagaaa agattaaaca gtgt	24
<210> 93 <211> 24 <212> DNA <213> Homo sapiens	
<400> 93 tatttgagaa aagattaaac agtg	24
<210> 94 <211> 24 <212> DNA <213> Homo sapiens	
<400> 94 atcttttctc aaataaaaag acat	24
<210> 95 <211> 24 <212> DNA <213> Homo sapiens	
<400> 95 ctcaaataaa aagacatggg gctt	24
<210> 96 <211> 24 <212> DNA <213> Homo sapiens	
<400> 96	

tcaaataaaa agacatgggg cttc 24

<210> 97
<211> 24
<212> DNA
<213> Homo sapiens

<400> 97
tgttttgcct ttttggatc ttac 24

<210> 98
<211> 24
<212> DNA
<213> Homo sapiens

<400> 98
gttttgcctt tttggatct taca 24

<210> 99
<211> 24
<212> DNA
<213> Homo sapiens

<400> 99
ttttgccttt ttggatctt acag 24

<210> 100
<211> 24
<212> DNA
<213> Homo sapiens

<400> 100
gcctttttgg tatcttacag gaac 24

<210> 101
<211> 24
<212> DNA
<213> Homo sapiens

<400> 101
cctttttggg atcttacagg aact 24

<210> 102
<211> 24
<212> DNA
<213> Homo sapiens

<400> 102
tggtatctta caggaactcc agga 24

<210> 103
<211> 24
<212> DNA
<213> Homo sapiens

<400> 103
ggtatcttac aggaactcca ggat 24

<210> 104	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 104	
aggattgctg aattatttct tccc	24
<210> 105	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 105	
gaggattgct gaattatttc ttcc	24
<210> 106	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 106	
tgaggattgc tgaattattt ctcc	24
<210> 107	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 107	
ctgtagaata ctggcatctg tttt	24
<210> 108	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 108	
cctgtagaat actggcatct gtgt	24
<210> 109	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 109	
tcctgtagaa tactggcatc tggt	24
<210> 110	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 110	
cagacctcct gccaccgcag attc	24

<210> 111	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 111	
tgtctgacag ctgtttgcag acct	24
<210> 112	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 112	
ctgtctgaca gctgtttgca gacc	24
<210> 113	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 113	
tctgtctgac agctgtttgc agac	24
<210> 114	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 114	
ttctgtctga cagctgtttg caga	24
<210> 115	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 115	
attcctatta gatctgtcgc ccta	24
<210> 116	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 116	
cattcctatt agatctgtcg ccct	24
<210> 117	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 117	
agcagacttt ttaagctttc tttta	24

<210> 118
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 118
 gcagactttt taagctttct ttag 24

 <210> 119
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 119
 taagctttct ttagaagaat attt 24

 <210> 120
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 120
 aagctttctt tagaagaata ttct 24

 <210> 121
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 121
 agctttcttt agaagaatat ttca 24

 <210> 122
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 122
 tttagaagaa tatttcatga gaga 24

 <210> 123
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 123
 gaagaatatt tcatgagaga ttat 24

 <210> 124
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 124
 tcagtataac caaaaaatat acgc 24

 <210> 125

<211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 125
 acataatcca tctatTTTTtC ttga 24

 <210> 126
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 126
 cataatccat ctatTTTTtCt tgat 24

 <210> 127
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 127
 tcttgatcca tatgctTTTTa cctg 24

 <210> 128
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 128
 cttgatccat atgctTTTTac ctgc 24

 <210> 129
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 129
 ttgatccata tgctTTTTacc tgca 24

 <210> 130
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 130
 tcaacagatc tgtcaaatcg cctg 24

 <210> 131
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 131
 acctgcaggc gatttgacag atct 24

 <210> 132
 <211> 24

<212> DNA
 <213> Homo sapiens

<400> 132
 cctgcaggcg atttgacaga tctg 24

<210> 133
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 133
 acagatctgt tgagaaatgg cggc 24

<210> 134
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 134
 tatcataatg aaaacgccgc catt 24

<210> 135
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 135
 cattatgata taaagatatt taat 24

<210> 136
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 136
 tatttagcat gttcccaatt ctca 24

<210> 137
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 137
 gaaaaaaca atcaaagact tacc 24

<210> 138
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 138
 atttgttttt tcgaaattgt attt 24

<210> 139
 <211> 24
 <212> DNA

<213> Homo sapiens

<400> 139

tttttttcgaa attgtattta tctt

24

<210> 140

<211> 24

<212> DNA

<213> Homo sapiens

<400> 140

ttcgaaattg tatttatctt cagc

24

<210> 141

<211> 24

<212> DNA

<213> Homo sapiens

<400> 141

tcgaaattgt atttatcttc agca

24

<210> 142

<211> 24

<212> DNA

<213> Homo sapiens

<400> 142

cgaaattgta tttatcttca gcac

24

<210> 143

<211> 24

<212> DNA

<213> Homo sapiens

<400> 143

gaaattgtat ttatcttcag caca

24

<210> 144

<211> 24

<212> DNA

<213> Homo sapiens

<400> 144

agaagttaaa gagtcagat gtgc

24

<210> 145

<211> 24

<212> DNA

<213> Homo sapiens

<400> 145

tcttcagcac atctggactc tttta

24

<210> 146

<211> 24

<212> DNA

<213> Homo sapiens

<400> 146
catcaccctt cagaacctga tctt 24

<210> 147
<211> 24
<212> DNA
<213> Homo sapiens

<400> 147
acttcttaaa gatcagggttc tgaa 24

<210> 148
<211> 24
<212> DNA
<213> Homo sapiens

<400> 148
gactggtggt gtcattatta tatt 24

<210> 149
<211> 24
<212> DNA
<213> Homo sapiens

<400> 149
actggtgttg tcatcattat atta 24

<210> 150
<211> 24
<212> DNA
<213> Homo sapiens

<400> 150
aactagaata aaaggaaaaa taaa 24

<210> 151
<211> 24
<212> DNA
<213> Homo sapiens

<400> 151
ctactatata tttatctttc cttt 24

<210> 152
<211> 24
<212> DNA
<213> Homo sapiens

<400> 152
tttttccttt tattctagtt gaaa 24

<210> 153
<211> 24
<212> DNA
<213> Homo sapiens

<400> 153 tccttttatt ctagttgaaa gaat	24
<210> 154 <211> 24 <212> DNA <213> Homo sapiens	
<400> 154 cctttttattc tagttgaaag aatt	24
<210> 155 <211> 24 <212> DNA <213> Homo sapiens	
<400> 155 ctttttattct agttgaaaga attc	24
<210> 156 <211> 24 <212> DNA <213> Homo sapiens	
<400> 156 attctagttg aaagaattca gaat	24
<210> 157 <211> 24 <212> DNA <213> Homo sapiens	
<400> 157 ttctagttga aagaattcag aatc	24
<210> 158 <211> 24 <212> DNA <213> Homo sapiens	
<400> 158 attcaactgt tgcctccggt tctg	24
<210> 159 <211> 24 <212> DNA <213> Homo sapiens	
<400> 159 acatttcatt caactggttc ctcc	24
<210> 160 <211> 24 <212> DNA <213> Homo sapiens	
<400> 160	

cttttggatt gcatctactg tata 24

<210> 161
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 161
 tgtgattttc ttttggattg catc 24

<210> 162
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 162
 atactaacct tggtttctgt gatt 24

<210> 163
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 163
 aaaaggtatc tttgatacta acct 24

<210> 164
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 164
 aaaaaggtat ctttgatact aacc 24

<210> 165
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 165
 ttttaaaaag gtatctttga tact 24

<210> 166
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 166
 attttaaaaa ggtatctttg atac 24

<210> 167
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 167
 ttaatgcaaa ctgggacaca aaca 24

<210> 168	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 168	
taaattgccca tgtttgtgtc ccag	24
<210> 169	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 169	
aaattgccat gtttgtgtcc cagt	24
<210> 170	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 170	
aattgccatg tttgtgtccc agtt	24
<210> 171	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 171	
tgtcccagtt tgcattaaca aata	24
<210> 172	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 172	
caacatagtt ctcaaactat ttgt	24
<210> 173	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 173	
ccaacatagt tctcaaacta ttgt	24
<210> 174	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 174	
tccaacatag ttctcaaact attt	24

<210> 175	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 175	
tttccaacat agttctcaaa ctat	24
<210> 176	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 176	
ttttccaaca tagttctcaa acta	24
<210> 177	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 177	
tttttccaac atagttctca aact	24
<210> 178	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 178	
cattaacaaa tagtttgaga acta	24
<210> 179	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 179	
agaactatgt tggaaaaaaaa aata	24
<210> 180	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 180	
gttcttctag cctggagaaa gaag	24
<210> 181	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 181	
attcttcttt ctccaggcta gaag	24

<210> 182
<211> 24
<212> DNA
<213> Homo sapiens

<400> 182
ttctttcttttc tccaggctag aaga

24

<210> 183
<211> 24
<212> DNA
<213> Homo sapiens

<400> 183
tccaggctag aagaacaaaa gaat

24

<210> 184
<211> 24
<212> DNA
<213> Homo sapiens

<400> 184
aaattctgac aagatattct ttg

24

<210> 185
<211> 24
<212> DNA
<213> Homo sapiens

<400> 185
cttttagttg ctgctctttt ccag

24

<210> 186
<211> 24
<212> DNA
<213> Homo sapiens

<400> 186
agaaaataaa attaccttga cttg

24

<210> 187
<211> 24
<212> DNA
<213> Homo sapiens

<400> 187
tgcaagcagg ccctggggga ttg

24

<210> 188
<211> 24
<212> DNA
<213> Homo sapiens

<400> 188
attttctcaa atcccccagg gcct

24

<210> 189

<211> 24
<212> DNA
<213> Homo sapiens

<400> 189
ttttctcaaa tccccaggg cctg 24

<210> 190
<211> 24
<212> DNA
<213> Homo sapiens

<400> 190
ctcaaattccc ccagggcctg cttg 24

<210> 191
<211> 24
<212> DNA
<213> Homo sapiens

<400> 191
tcaaattcccc cagggcctgc ttgc 24

<210> 192
<211> 24
<212> DNA
<213> Homo sapiens

<400> 192
ttaattcaat cattggtttt ctgc 24

<210> 193
<211> 24
<212> DNA
<213> Homo sapiens

<400> 193
taattcaatc attggttttc tgcc 24

<210> 194
<211> 24
<212> DNA
<213> Homo sapiens

<400> 194
aattcaatca ttggttttct gccc 24

<210> 195
<211> 24
<212> DNA
<213> Homo sapiens

<400> 195
attcaatcat tggttttctg ccca 24

<210> 196
<211> 24

<212> DNA	
<213> Homo sapiens	
<400> 196	
gcaaggaact atgaataacc taat	24
<210> 197	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 197	
ctgcccatta gggtattcat agtt	24
<210> 198	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 198	
tgcccattag gttattcata gttc	24
<210> 199	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 199	
tagaaaacaa tttaacagga aata	24
<210> 200	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 200	
ctgttaaatt gttttctata aacc	24
<210> 201	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 201	
gaaataaaaa agatgttact gtat	24
<210> 202	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 202	
agaaataaaa aagatgttac tgta	24
<210> 203	
<211> 24	
<212> DNA	

<213> Homo sapiens

<400> 203
ctataaaccc ttatacagta acat 24

<210> 204
<211> 24
<212> DNA
<213> Homo sapiens

<400> 204
tataaacct tatacagtaa catc 24

<210> 205
<211> 24
<212> DNA
<213> Homo sapiens

<400> 205
ttatttctaa aagtgttttg gctg 24

<210> 206
<211> 24
<212> DNA
<213> Homo sapiens

<400> 206
tatttctaaa agtgttttgg ctgg 24

<210> 207
<211> 24
<212> DNA
<213> Homo sapiens

<400> 207
atttctaataa gtgttttggc tggc 24

<210> 208
<211> 24
<212> DNA
<213> Homo sapiens

<400> 208
tttctaataa tggttttggt ggct 24

<210> 209
<211> 24
<212> DNA
<213> Homo sapiens

<400> 209
taaaagtgtt ttggctggct tcac 24

<210> 210
<211> 24
<212> DNA
<213> Homo sapiens

<400> 210 cataatacaa agtaaagtac aatt	24
<210> 211 <211> 24 <212> DNA <213> Homo sapiens	
<400> 211 acataataca aagtaaagta caat	24
<210> 212 <211> 24 <212> DNA <213> Homo sapiens	
<400> 212 ggctgggtctc acaattgtac tttta	24
<210> 213 <211> 24 <212> DNA <213> Homo sapiens	
<400> 213 gctgggtctca caattgtact ttac	24
<210> 214 <211> 24 <212> DNA <213> Homo sapiens	
<400> 214 ctttgtatta tgtaaaagga atac	24
<210> 215 <211> 24 <212> DNA <213> Homo sapiens	
<400> 215 tattatgtaa aaggaataca caac	24
<210> 216 <211> 24 <212> DNA <213> Homo sapiens	
<400> 216 ttcttacagg caacaatgca ggat	24
<210> 217 <211> 24 <212> DNA <213> Homo sapiens	

<400> 217
 gaacagaggc gtccccagtt ggaa 24

<210> 218
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 218
 ggcagcggta atgagttctt ccaa 24

<210> 219
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 219
 tcaaattttg ggcagcggta atga 24

<210> 220
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 220
 aaaaacaaga ccagcaatca agag 24

<210> 221
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 221
 tgtgtcccat gcttggttaa aaac 24

<210> 222
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 222
 ttaacaagca tgggacacac aaag 24

<210> 223
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 223
 taacaagcat gggacacaca aagc 24

<210> 224
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 224

aacaagcatg ggacacacaa agca 24

<210> 225
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 225
 acaagcatgg gacacacaaa gcaa 24

<210> 226
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 226
 ttgaaacttg tcatgcatct tgct 24

<210> 227
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 227
 attgaaactt gtcatgcatc ttgc 24

<210> 228
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 228
 tattgaaact tgatcatgat cttg 24

<210> 229
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 229
 aataaaaact taagttcata tadc 24

<210> 230
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 230
 gtgaatatat tattggattt ctat 24

<210> 231
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 231
 aagataattc atgaacatct taat 24

<210> 232	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 232	
acagaaaagc atacacatta cttta	24
<210> 233	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 233	
ctgttaaaga ggaagttaga agat	24
<210> 234	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 234	
tgttaaagag gaagttagaa gatc	24
<210> 235	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 235	
ccgccttcca ctcagagctc agat	24
<210> 236	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 236	
ccctcagctc ttgaagtaaa cggt	24
<210> 237	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 237	
cttcaagagc tgagggcaaa gcag	24
<210> 238	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 238	
aacaaatagc tagagccaaa gaga	24

<210> 239	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 239	
gaacaaatag ctagagccaa agag	24
<210> 240	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 240	
gctctagcta tttgttcaaa agtg	24
<210> 241	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 241	
ttcaaaagtg caactatgaa gtga	24
<210> 242	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 242	
tctctcacc agtcatcact tcat	24
<210> 243	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 243	
ctctctcacc cagtcacac ttca	24
<210> 244	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 244	
tatatatata tatatTTTTc tctt	24
<210> 245	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 245	
tctctttcta tagacagcta attc	24

<210> 246
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 246
 ctcttttctat agacagctaa ttca 24

 <210> 247
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 247
 aaacagtaaa aaaatgaatt agct 24

 <210> 248
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 248
 tcttttctata gacagctaatt tcat 24

 <210> 249
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 249
 aaaacagtaa aaaaatgaat tagc 24

 <210> 250
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 250
 tatagacagc taattcatttt tttt 24

 <210> 251
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 251
 tatttctgtaa tataaaaaatt ttaa 24

 <210> 252
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 252
 atattctgta atataaaaaat ttta 24

 <210> 253

<211> 24
<212> DNA
<213> Homo sapiens

<400> 253
tttactgttt taaaattttt atat 24

<210> 254
<211> 24
<212> DNA
<213> Homo sapiens

<400> 254
ttactgtttt aaaattttta tatt 24

<210> 255
<211> 24
<212> DNA
<213> Homo sapiens

<400> 255
tactgtttta aaatttttat atta 24

<210> 256
<211> 24
<212> DNA
<213> Homo sapiens

<400> 256
actgttttaa aatttttata ttac 24

<210> 257
<211> 24
<212> DNA
<213> Homo sapiens

<400> 257
ctgtttttaa atttttatat taca 24

<210> 258
<211> 24
<212> DNA
<213> Homo sapiens

<400> 258
aaaattttta tattacagaa tata 24

<210> 259
<211> 24
<212> DNA
<213> Homo sapiens

<400> 259
aaatttttat attacagaat ataa 24

<210> 260
<211> 24

<212> DNA
 <213> Homo sapiens

<400> 260
 ttgtagacta tcttttatat tctg 24

<210> 261
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 261
 tatattacag aatataaaag atag 24

<210> 262
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 262
 atattacaga atataaaaga tagt 24

<210> 263
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 263
 tattacagaa tataaaagat agtc 24

<210> 264
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 264
 caatgctgct gtcttcttgc tatg 24

<210> 265
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 265
 caatgggaaa aagttaacaa aatg 24

<210> 266
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 266
 tgcaagtatc aagaaaaata tatg 24

<210> 267
 <211> 24
 <212> DNA

<213> Homo sapiens

<400> 267
tcttgatact tgcagaaatg attt 24

<210> 268
<211> 24
<212> DNA
<213> Homo sapiens

<400> 268
cttgatactt gcagaaatga ttgt 24

<210> 269
<211> 24
<212> DNA
<213> Homo sapiens

<400> 269
ttgatacttg cagaaatgat ttgt 24

<210> 270
<211> 24
<212> DNA
<213> Homo sapiens

<400> 270
ttttcagggga actgtagaat ttat 24

<210> 271
<211> 24
<212> DNA
<213> Homo sapiens

<400> 271
catggaggggt actgaaataa attc 24

<210> 272
<211> 24
<212> DNA
<213> Homo sapiens

<400> 272
ccatggaggg tactgaaata aatt 24

<210> 273
<211> 24
<212> DNA
<213> Homo sapiens

<400> 273
caggggaactg tagaatttat ttca 24

<210> 274
<211> 24
<212> DNA
<213> Homo sapiens

<400> 274 tccatggagg gtactgaaat aaat	24
<210> 275 <211> 24 <212> DNA <213> Homo sapiens	
<400> 275 agggaactgt agaatttatt tcag	24
<210> 276 <211> 24 <212> DNA <213> Homo sapiens	
<400> 276 ttccatggag ggtactgaaa taaa	24
<210> 277 <211> 24 <212> DNA <213> Homo sapiens	
<400> 277 cctgtctttt ttccatggag ggta	24
<210> 278 <211> 24 <212> DNA <213> Homo sapiens	
<400> 278 ccctgtcttt tttccatgga ggggt	24
<210> 279 <211> 24 <212> DNA <213> Homo sapiens	
<400> 279 tccctgtctt ttttccatgg aggg	24
<210> 280 <211> 24 <212> DNA <213> Homo sapiens	
<400> 280 tttcagtagc ctccatggaa aaaa	24
<210> 281 <211> 24 <212> DNA <213> Homo sapiens	

<400> 281
agtaccctcc atggaaaaaa gaca 24

<210> 282
<211> 24
<212> DNA
<213> Homo sapiens

<400> 282
agtttgcattg gttcttgctc aagg 24

<210> 283
<211> 24
<212> DNA
<213> Homo sapiens

<400> 283
ataagaaaat gcattccttg agca 24

<210> 284
<211> 24
<212> DNA
<213> Homo sapiens

<400> 284
cataagaaaa tgcattcctt gagc 24

<210> 285
<211> 24
<212> DNA
<213> Homo sapiens

<400> 285
catggttctt gctcaaggaa tgca 24

<210> 286
<211> 24
<212> DNA
<213> Homo sapiens

<400> 286
acctacatgt ggaaataaat tttc 24

<210> 287
<211> 24
<212> DNA
<213> Homo sapiens

<400> 287
gacctacatg tggaataaaa tttt 24

<210> 288
<211> 24
<212> DNA
<213> Homo sapiens

<400> 288

tgacctacat gtggaaataa attt 24

<210> 289

<211> 24

<212> DNA

<213> Homo sapiens

<400> 289

cttatgaaaa tttattttcca catg 24

<210> 290

<211> 24

<212> DNA

<213> Homo sapiens

<400> 290

ttatgaaaat ttattttccac atgt 24

<210> 291

<211> 24

<212> DNA

<213> Homo sapiens

<400> 291

attacatttt tgacctacat gtgg 24

<210> 292

<211> 24

<212> DNA

<213> Homo sapiens

<400> 292

cattacattt ttgacctaca tgtg 24

<210> 293

<211> 24

<212> DNA

<213> Homo sapiens

<400> 293

tcattacatt tttgacctac atgt 24

<210> 294

<211> 24

<212> DNA

<213> Homo sapiens

<400> 294

tttccacatg taggtcaaaa atgt 24

<210> 295

<211> 24

<212> DNA

<213> Homo sapiens

<400> 295

cacatgtagg tcaaaaatgt aatg 24

<210> 296	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 296	
ttgcaatcca gccatgatat tttt	24
<210> 297	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 297	
actggttggtt tggtgcaatc cagc	24
<210> 298	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 298	
cactgttggt ttgttgcaat ccag	24
<210> 299	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 299	
aatgctctca tccatagtca tagg	24
<210> 300	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 300	
atgtctcagt aatcttctta ccta	24
<210> 301	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 301	
caagttattt aatgtctcag taat	24
<210> 302	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 302	
acaagttatt taatgtctca gtaa	24

<210> 303	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 303	
gactctgatg acatattttt cccc	24
<210> 304	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 304	
tccccagtat ggttccagat catg	24
<210> 305	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 305	
ccccagtatg gttccagatc atgt	24
<210> 306	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 306	
cccagtatgg ttccagatca tgtc	24
<210> 307	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 307	
tatttgtctt tgtgtatgtg tgta	24
<210> 308	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 308	
tctttgtgta tgtgtgtatg tgta	24
<210> 309	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 309	
tgtatgtgtg tatgtgtatg tggt	24

<210> 310
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 310
 aggccagacc tatttgactg gaat 24

 <210> 311
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 311
 ggccagacct atttgactgg aata 24

 <210> 312
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 312
 actggaatag tgtggtttgc cagc 24

 <210> 313
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 313
 ccagcagtca gccacacaac gact 24

 <210> 314
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 314
 tctatgccta attgatattct ggcg 24

 <210> 315
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 315
 ccaaccttca ggatcgagta gttt 24

 <210> 316
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 316
 tggactacca ctgcttttag tatg 24

 <210> 317

<211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 317
 agtatggtag agtttaatgt tttc 24

 <210> 318
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 318
 gtatggtaga gtttaatggt ttca 24

 <210> 319
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 319
 agactctaaa aggataatga acaa 24

 <210> 320
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 320
 actttgattt gttcattatc cttt 24

 <210> 321
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 321
 tatatttgag actctaaaag gata 24

 <210> 322
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 322
 atttgttcat tatectttta gagt 24

 <210> 323
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 323
 gtttctatat ttgagactct aaaa 24

 <210> 324
 <211> 24

<212> DNA	
<213> Homo sapiens	
<400> 324	
ggtttctata tttgagactc taaa	24
<210> 325	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 325	
tggtttctat atttgagact ctaa	24
<210> 326	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 326	
ttcattatcc ttttagagtc tcaa	24
<210> 327	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 327	
agagtctcaa atatagaaac caaa	24
<210> 328	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 328	
gagtctcaaa tatagaaacc aaaa	24
<210> 329	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 329	
cacttctctgg atggcttcaa tgct	24
<210> 330	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 330	
gcctcaacaa gtgagcattg aagc	24
<210> 331	
<211> 24	
<212> DNA	

<213> Homo sapiens	
<400> 331	
cctcaacaag tgagcattga agcc	24
<210> 332	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 332	
ggtggccttg gcaacatttc cact	24
<210> 333	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 333	
gtcacttttag gtggccttgg caac	24
<210> 334	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 334	
atgatgtaac tgaaaatggt cttc	24
<210> 335	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 335	
cctggttgaga atagtgcatt tgat	24
<210> 336	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 336	
cagttacatc atcaaattgca ctat	24
<210> 337	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 337	
agttacatca tcaaattgcac tatt	24
<210> 338	
<211> 24	
<212> DNA	
<213> Homo sapiens	

<400> 338 cacactttac ctgttgagaa tagt	24
<210> 339 <211> 24 <212> DNA <213> Homo sapiens	
<400> 339 ctgttttata tgcattttta ggta	24
<210> 340 <211> 24 <212> DNA <213> Homo sapiens	
<400> 340 tgttttatat gcatttttag gtat	24
<210> 341 <211> 24 <212> DNA <213> Homo sapiens	
<400> 341 atatgcattt ttaggtatta cgtg	24
<210> 342 <211> 24 <212> DNA <213> Homo sapiens	
<400> 342 tatgcatttt taggtattac gtgc	24
<210> 343 <211> 24 <212> DNA <213> Homo sapiens	
<400> 343 taggtattac gtgcacatat atat	24
<210> 344 <211> 24 <212> DNA <213> Homo sapiens	
<400> 344 aggtattacg tgcacatata tata	24
<210> 345 <211> 24 <212> DNA <213> Homo sapiens	

<400> 345 ggtattacgt gcacatatat atat	24
<210> 346 <211> 24 <212> DNA <213> Homo sapiens	
<400> 346 agcaacaact ataatatattgt gcag	24
<210> 347 <211> 24 <212> DNA <213> Homo sapiens	
<400> 347 gttcctccat ctttctcttt ttat	24
<210> 348 <211> 24 <212> DNA <213> Homo sapiens	
<400> 348 tctttttatg gagttcacta ggtg	24
<210> 349 <211> 24 <212> DNA <213> Homo sapiens	
<400> 349 tatggagttc actaggtgca ccat	24
<210> 350 <211> 24 <212> DNA <213> Homo sapiens	
<400> 350 atggagttca ctaggtgcac catt	24
<210> 351 <211> 24 <212> DNA <213> Homo sapiens	
<400> 351 tggagttcac taggtgcacc attc	24
<210> 352 <211> 24 <212> DNA <213> Homo sapiens	
<400> 352	

ataattgcat ctgaacattt ggtc 24

<210> 353
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 353
 gtcctttgca gggtagagtga gcga 24

<210> 354
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 354
 ttccaaagca gcctctcgct cact 24

<210> 355
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 355
 cagggtagagt gagcgagagg ctgc 24

<210> 356
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 356
 gaagaaactc atagattact gcaa 24

<210> 357
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 357
 caggtccagg gggaactggt gcag 24

<210> 358
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 358
 ccaggtccag ggggaactgt tgca 24

<210> 359
 <211> 24
 <212> DNA
 <213> Homo sapiens

<400> 359
 agcttctgta agccaggcaa gaaa 24

<210> 360	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 360	
ttgcctggct tacagaagct gaaa	24
<210> 361	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 361	
cttacgggta gcatcctgta ggac	24
<210> 362	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 362	
ctcccttgga gtcttctagg agcc	24
<210> 363	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 363	
actcccttgg agtcttctag gagc	24
<210> 364	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 364	
atcagctctt ttactccctt ggag	24
<210> 365	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 365	
cgctttagca ctcttggtga tcca	24
<210> 366	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 366	
gcactcttgt ggatccaatt gaac	24

<210> 367	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 367	
tccctggcctt gtcagttaca agta	24
<210> 368	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 368	
gtccctggct tgtcagttac aagt	24
<210> 369	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 369	
ttttgtccct ggcttgtcag ttac	24
<210> 370	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 370	
gttttgtccc tggcttgtca gtta	24
<210> 371	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 371	
tacttgtaac tgacaagcca ggga	24
<210> 372	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 372	
gctcctactc agactgttac tctg	24
<210> 373	
<211> 24	
<212> DNA	
<213> Homo sapiens	
<400> 373	
taccatgtat tgctaaacaa agta	24

<210> 374
 <211> 24
 <212> DNA
 <213> Homo sapiens

 <400> 374
 attgaagagt aacaatttga gcca 24

 <210> 375
 <211> 24
 <212> DNA
 <213> Mus sp.

 <400> 375
 aggctctgca aagttctttg aaag 24

 <210> 376
 <211> 24
 <212> DNA
 <213> Mus sp.

 <400> 376
 aaagagcaac aaaatggctt caac 24

 <210> 377
 <211> 24
 <212> DNA
 <213> Mus sp.

 <400> 377
 aaagagcaat aaaatggctt caac 24

 <210> 378
 <211> 24
 <212> DNA
 <213> Mus sp.

 <400> 378
 aaagaacttt gcagagcctc aaaa 24

 <210> 379
 <211> 24
 <212> DNA
 <213> Mus sp.

 <400> 379
 ctgaatatct atgcattaat aact 24

 <210> 380
 <211> 24
 <212> DNA
 <213> Mus sp.

 <400> 380
 tattatatta cagggcataat tata 24

 <210> 381

<211> 24
 <212> DNA
 <213> Mus sp.

<400> 381
 aggtaagccg aggtttggcc tttta

24

<210> 382
 <211> 24
 <212> DNA
 <213> Mus sp.

<400> 382
 cccagagtcc ttcaaagata ttga

24

<210> 383
 <211> 3685
 <212> PRT
 <213> Homo sapiens

<400> 383
 Met Leu Trp Trp Glu Glu Val Glu Asp Cys Tyr Glu Arg Glu Asp Val
 1 5 10 15

Gln Lys Lys Thr Phe Thr Lys Trp Val Asn Ala Gln Phe Ser Lys Phe
 20 25 30

Gly Lys Gln His Ile Glu Asn Leu Phe Ser Asp Leu Gln Asp Gly Arg
 35 40 45

Arg Leu Leu Asp Leu Leu Glu Gly Leu Thr Gly Gln Lys Leu Pro Lys
 50 55 60

Glu Lys Gly Ser Thr Arg Val His Ala Leu Asn Asn Val Asn Lys Ala
 65 70 75 80

Leu Arg Val Leu Gln Asn Asn Asn Val Asp Leu Val Asn Ile Gly Ser
 85 90 95

Thr Asp Ile Val Asp Gly Asn His Lys Leu Thr Leu Gly Leu Ile Trp
 100 105 110

Asn Ile Ile Leu His Trp Gln Val Lys Asn Val Met Lys Asn Ile Met
 115 120 125

Ala Gly Leu Gln Gln Thr Asn Ser Glu Lys Ile Leu Leu Ser Trp Val
 130 135 140

Arg Gln Ser Thr Arg Asn Tyr Pro Gln Val Asn Val Ile Asn Phe Thr
 145 150 155 160

Thr Ser Trp Ser Asp Gly Leu Ala Leu Asn Ala Leu Ile His Ser His

165	170	175
Arg Pro Asp Leu Phe Asp Trp Asn Ser Val Val Cys Gln Gln Ser Ala		
180	185	190
Thr Gln Arg Leu Glu His Ala Phe Asn Ile Ala Arg Tyr Gln Leu Gly		
195	200	205
Ile Glu Lys Leu Leu Asp Pro Glu Asp Val Asp Thr Thr Tyr Pro Asp		
210	215	220
Lys Lys Ser Ile Leu Met Tyr Ile Thr Ser Leu Phe Gln Val Leu Pro		
225	230	235
Gln Gln Val Ser Ile Glu Ala Ile Gln Glu Val Glu Met Leu Pro Arg		
245	250	255
Pro Pro Lys Val Thr Lys Glu Glu His Phe Gln Leu His His Gln Met		
260	265	270
His Tyr Ser Gln Gln Ile Thr Val Ser Leu Ala Gln Gly Tyr Glu Arg		
275	280	285
Thr Ser Ser Pro Lys Pro Arg Phe Lys Ser Tyr Ala Tyr Thr Gln Ala		
290	295	300
Ala Tyr Val Thr Thr Ser Asp Pro Thr Arg Ser Pro Phe Pro Ser Gln		
305	310	315
His Leu Glu Ala Pro Glu Asp Lys Ser Phe Gly Ser Ser Leu Met Glu		
325	330	335
Ser Glu Val Asn Leu Asp Arg Tyr Gln Thr Ala Leu Glu Glu Val Leu		
340	345	350
Ser Trp Leu Leu Ser Ala Glu Asp Thr Leu Gln Ala Gln Gly Glu Ile		
355	360	365
Ser Asn Asp Val Glu Val Val Lys Asp Gln Phe His Thr His Glu Gly		
370	375	380
Tyr Met Met Asp Leu Thr Ala His Gln Gly Arg Val Gly Asn Ile Leu		
385	390	395
Gln Leu Gly Ser Lys Leu Ile Gly Thr Gly Lys Leu Ser Glu Asp Glu		
405	410	415
Glu Thr Glu Val Gln Glu Gln Met Asn Leu Leu Asn Ser Arg Trp Glu		

420	425	430																	
Cys	Leu	Arg	Val	Ala	Ser	Met	Glu	Lys	Gln	Ser	Asn	Leu	His	Arg	Val				
	435						440					445							
Leu	Met	Asp	Leu	Gln	Asn	Gln	Lys	Leu	Lys	Glu	Leu	Asn	Asp	Trp	Leu				
	450					455					460								
Thr	Lys	Thr	Glu	Glu	Arg	Thr	Arg	Lys	Met	Glu	Glu	Glu	Pro	Leu	Gly				
465					470					475					480				
Pro	Asp	Leu	Glu	Asp	Leu	Lys	Arg	Gln	Val	Gln	Gln	His	Lys	Val	Leu				
				485					490					495					
Gln	Glu	Asp	Leu	Glu	Gln	Glu	Gln	Val	Arg	Val	Asn	Ser	Leu	Thr	His				
			500					505					510						
Met	Val	Val	Val	Val	Asp	Glu	Ser	Ser	Gly	Asp	His	Ala	Thr	Ala	Ala				
		515					520					525							
Leu	Glu	Glu	Gln	Leu	Lys	Val	Leu	Gly	Asp	Arg	Trp	Ala	Asn	Ile	Cys				
	530					535					540								
Arg	Trp	Thr	Glu	Asp	Arg	Trp	Val	Leu	Leu	Gln	Asp	Ile	Leu	Leu	Lys				
545					550					555					560				
Trp	Gln	Arg	Leu	Thr	Glu	Glu	Gln	Cys	Leu	Phe	Ser	Ala	Trp	Leu	Ser				
				565					570					575					
Glu	Lys	Glu	Asp	Ala	Val	Asn	Lys	Ile	His	Thr	Thr	Gly	Phe	Lys	Asp				
			580					585					590						
Gln	Asn	Glu	Met	Leu	Ser	Ser	Leu	Gln	Lys	Leu	Ala	Val	Leu	Lys	Ala				
		595					600					605							
Asp	Leu	Glu	Lys	Lys	Lys	Gln	Ser	Met	Gly	Lys	Leu	Tyr	Ser	Leu	Lys				
	610					615					620								
Gln	Asp	Leu	Leu	Ser	Thr	Leu	Lys	Asn	Lys	Ser	Val	Thr	Gln	Lys	Thr				
625					630					635					640				
Glu	Ala	Trp	Leu	Asp	Asn	Phe	Ala	Arg	Cys	Trp	Asp	Asn	Leu	Val	Gln				
				645					650					655					
Lys	Leu	Glu	Lys	Ser	Thr	Ala	Gln	Ile	Ser	Gln	Ala	Val	Thr	Thr	Thr				
			660					665					670						
Gln	Pro	Ser	Leu	Thr	Gln	Thr	Thr	Val	Met	Glu	Thr	Val	Thr	Thr	Val				

675	680	685																	
Thr	Thr	Arg	Glu	Gln	Ile	Leu	Val	Lys	His	Ala	Gln	Glu	Glu	Leu	Pro				
690						695					700								
Pro	Pro	Pro	Pro	Gln	Lys	Lys	Arg	Gln	Ile	Thr	Val	Asp	Ser	Glu	Ile				
705					710					715					720				
Arg	Lys	Arg	Leu	Asp	Val	Asp	Ile	Thr	Glu	Leu	His	Ser	Trp	Ile	Thr				
				725					730					735					
Arg	Ser	Glu	Ala	Val	Leu	Gln	Ser	Pro	Glu	Phe	Ala	Ile	Phe	Arg	Lys				
			740					745					750						
Glu	Gly	Asn	Phe	Ser	Asp	Leu	Lys	Glu	Lys	Val	Asn	Ala	Ile	Glu	Arg				
		755					760					765							
Glu	Lys	Ala	Glu	Lys	Phe	Arg	Lys	Leu	Gln	Asp	Ala	Ser	Arg	Ser	Ala				
	770					775					780								
Gln	Ala	Leu	Val	Glu	Gln	Met	Val	Asn	Glu	Gly	Val	Asn	Ala	Asp	Ser				
785					790					795					800				
Ile	Lys	Gln	Ala	Ser	Glu	Gln	Leu	Asn	Ser	Arg	Trp	Ile	Glu	Phe	Cys				
				805					810					815					
Gln	Leu	Leu	Ser	Glu	Arg	Leu	Asn	Trp	Leu	Glu	Tyr	Gln	Asn	Asn	Ile				
			820					825					830						
Ile	Ala	Phe	Tyr	Asn	Gln	Leu	Gln	Gln	Leu	Glu	Gln	Met	Thr	Thr	Thr				
	835						840					845							
Ala	Glu	Asn	Trp	Leu	Lys	Ile	Gln	Pro	Thr	Thr	Pro	Ser	Glu	Pro	Thr				
	850					855					860								
Ala	Ile	Lys	Ser	Gln	Leu	Lys	Ile	Cys	Lys	Asp	Glu	Val	Asn	Arg	Leu				
865					870					875					880				
Ser	Gly	Leu	Gln	Pro	Gln	Ile	Glu	Arg	Leu	Lys	Ile	Gln	Ser	Ile	Ala				
				885					890					895					
Leu	Lys	Glu	Lys	Gly	Gln	Gly	Pro	Met	Phe	Leu	Asp	Ala	Asp	Phe	Val				
			900					905					910						
Ala	Phe	Thr	Asn	His	Phe	Lys	Gln	Val	Phe	Ser	Asp	Val	Gln	Ala	Arg				
	915						920					925							
Glu	Lys	Glu	Leu	Gln	Thr	Ile	Phe	Asp	Thr	Leu	Pro	Pro	Met	Arg	Tyr				

1175		1180		1185
Thr Pro Asp Glu Leu Gln Lys Ala Val Glu Glu Met Lys Arg Ala				
1190		1195		1200
Lys Glu Glu Ala Gln Gln Lys Glu Ala Lys Val Lys Leu Leu Thr				
1205		1210		1215
Glu Ser Val Asn Ser Val Ile Ala Gln Ala Pro Pro Val Ala Gln				
1220		1225		1230
Glu Ala Leu Lys Lys Glu Leu Glu Thr Leu Thr Thr Asn Tyr Gln				
1235		1240		1245
Trp Leu Cys Thr Arg Leu Asn Gly Lys Cys Lys Thr Leu Glu Glu				
1250		1255		1260
Val Trp Ala Cys Trp His Glu Leu Leu Ser Tyr Leu Glu Lys Ala				
1265		1270		1275
Asn Lys Trp Leu Asn Glu Val Glu Phe Lys Leu Lys Thr Thr Glu				
1280		1285		1290
Asn Ile Pro Gly Gly Ala Glu Glu Ile Ser Glu Val Leu Asp Ser				
1295		1300		1305
Leu Glu Asn Leu Met Arg His Ser Glu Asp Asn Pro Asn Gln Ile				
1310		1315		1320
Arg Ile Leu Ala Gln Thr Leu Thr Asp Gly Gly Val Met Asp Glu				
1325		1330		1335
Leu Ile Asn Glu Glu Leu Glu Thr Phe Asn Ser Arg Trp Arg Glu				
1340		1345		1350
Leu His Glu Glu Ala Val Arg Arg Gln Lys Leu Leu Glu Gln Ser				
1355		1360		1365
Ile Gln Ser Ala Gln Glu Thr Glu Lys Ser Leu His Leu Ile Gln				
1370		1375		1380
Glu Ser Leu Thr Phe Ile Asp Lys Gln Leu Ala Ala Tyr Ile Ala				
1385		1390		1395
Asp Lys Val Asp Ala Ala Gln Met Pro Gln Glu Ala Gln Lys Ile				
1400		1405		1410
Gln Ser Asp Leu Thr Ser His Glu Ile Ser Leu Glu Glu Met Lys				

1415		1420		1425
Lys His Asn Gln Gly Lys Glu Ala Ala Gln Arg Val Leu Ser Gln				
1430		1435		1440
Ile Asp Val Ala Gln Lys Lys Leu Gln Asp Val Ser Met Lys Phe				
1445		1450		1455
Arg Leu Phe Gln Lys Pro Ala Asn Phe Glu Leu Arg Leu Gln Glu				
1460		1465		1470
Ser Lys Met Ile Leu Asp Glu Val Lys Met His Leu Pro Ala Leu				
1475		1480		1485
Glu Thr Lys Ser Val Glu Gln Glu Val Val Gln Ser Gln Leu Asn				
1490		1495		1500
His Cys Val Asn Leu Tyr Lys Ser Leu Ser Glu Val Lys Ser Glu				
1505		1510		1515
Val Glu Met Val Ile Lys Thr Gly Arg Gln Ile Val Gln Lys Lys				
1520		1525		1530
Gln Thr Glu Asn Pro Lys Glu Leu Asp Glu Arg Val Thr Ala Leu				
1535		1540		1545
Lys Leu His Tyr Asn Glu Leu Gly Ala Lys Val Thr Glu Arg Lys				
1550		1555		1560
Gln Gln Leu Glu Lys Cys Leu Lys Leu Ser Arg Lys Met Arg Lys				
1565		1570		1575
Glu Met Asn Val Leu Thr Glu Trp Leu Ala Ala Thr Asp Met Glu				
1580		1585		1590
Leu Thr Lys Arg Ser Ala Val Glu Gly Met Pro Ser Asn Leu Asp				
1595		1600		1605
Ser Glu Val Ala Trp Gly Lys Ala Thr Gln Lys Glu Ile Glu Lys				
1610		1615		1620
Gln Lys Val His Leu Lys Ser Ile Thr Glu Val Gly Glu Ala Leu				
1625		1630		1635
Lys Thr Val Leu Gly Lys Lys Glu Thr Leu Val Glu Asp Lys Leu				
1640		1645		1650
Ser Leu Leu Asn Ser Asn Trp Ile Ala Val Thr Ser Arg Ala Glu				

1655		1660		1665
Glu Trp Leu Asn Leu Leu Leu	Glu Tyr Gln Lys His	Met Glu Thr		
1670	1675	1680		
Phe Asp Gln Asn Val Asp His	Ile Thr Lys Trp Ile	Ile Gln Ala		
1685	1690	1695		
Asp Thr Leu Leu Asp Glu Ser	Glu Lys Lys Lys Pro	Gln Gln Lys		
1700	1705	1710		
Glu Asp Val Leu Lys Arg Leu	Lys Ala Glu Leu Asn	Asp Ile Arg		
1715	1720	1725		
Pro Lys Val Asp Ser Thr Arg	Asp Gln Ala Ala Asn	Leu Met Ala		
1730	1735	1740		
Asn Arg Gly Asp His Cys Arg	Lys Leu Val Glu Pro	Gln Ile Ser		
1745	1750	1755		
Glu Leu Asn His Arg Phe Ala	Ala Ile Ser His Arg	Ile Lys Thr		
1760	1765	1770		
Gly Lys Ala Ser Ile Pro Leu	Lys Glu Leu Glu Gln	Phe Asn Ser		
1775	1780	1785		
Asp Ile Gln Lys Leu Leu Glu	Pro Leu Glu Ala Glu	Ile Gln Gln		
1790	1795	1800		
Gly Val Asn Leu Lys Glu Glu	Asp Phe Asn Lys Asp	Met Asn Glu		
1805	1810	1815		
Asp Asn Glu Gly Thr Val Lys	Glu Leu Leu Gln Arg	Gly Asp Asn		
1820	1825	1830		
Leu Gln Gln Arg Ile Thr Asp	Glu Arg Lys Arg Glu	Glu Ile Lys		
1835	1840	1845		
Ile Lys Gln Gln Leu Leu Gln	Thr Lys His Asn Ala	Leu Lys Asp		
1850	1855	1860		
Leu Arg Ser Gln Arg Arg Lys	Lys Ala Leu Glu Ile	Ser His Gln		
1865	1870	1875		
Trp Tyr Gln Tyr Lys Arg Gln	Ala Asp Asp Leu Leu	Lys Cys Leu		
1880	1885	1890		
Asp Asp Ile Glu Lys Lys Leu	Ala Ser Leu Pro Glu	Pro Arg Asp		

1895	1900	1905
Glu Arg Lys Ile Lys Glu 1910	Ile Asp Arg Glu Leu 1915	Gln Lys Lys Lys 1920
Glu Glu Leu Asn Ala Val 1925	Arg Arg Gln Ala Glu 1930	Gly Leu Ser Glu 1935
Asp Gly Ala Ala Met Ala 1940	Val Glu Pro Thr Gln 1945	Ile Gln Leu Ser 1950
Lys Arg Trp Arg Glu Ile 1955	Glu Ser Lys Phe Ala 1960	Gln Phe Arg Arg 1965
Leu Asn Phe Ala Gln Ile 1970	His Thr Val Arg Glu 1975	Glu Thr Met Met 1980
Val Met Thr Glu Asp Met 1985	Pro Leu Glu Ile Ser 1990	Tyr Val Pro Ser 1995
Thr Tyr Leu Thr Glu Ile 2000	Thr His Val Ser Gln 2005	Ala Leu Leu Glu 2010
Val Glu Gln Leu Leu Asn 2015	Ala Pro Asp Leu Cys 2020	Ala Lys Asp Phe 2025
Glu Asp Leu Phe Lys Gln 2030	Glu Glu Ser Leu Lys 2035	Asn Ile Lys Asp 2040
Ser Leu Gln Gln Ser Ser 2045	Gly Arg Ile Asp Ile 2050	Ile His Ser Lys 2055
Lys Thr Ala Ala Leu Gln 2060	Ser Ala Thr Pro Val 2065	Glu Arg Val Lys 2070
Leu Gln Glu Ala Leu Ser 2075	Gln Leu Asp Phe Gln 2080	Trp Glu Lys Val 2085
Asn Lys Met Tyr Lys Asp 2090	Arg Gln Gly Arg Phe 2095	Asp Arg Ser Val 2100
Glu Lys Trp Arg Arg Phe 2105	His Tyr Asp Ile Lys 2110	Ile Phe Asn Gln 2115
Trp Leu Thr Glu Ala Glu 2120	Gln Phe Leu Arg Lys 2125	Thr Gln Ile Pro 2130
Glu Asn Trp Glu His Ala Lys	Tyr Lys Trp Tyr Leu	Lys Glu Leu

2135		2140		2145
Gln Asp 2150	Gly Ile Gly Gln	Arg 2155	Gln Thr Val Val	Arg Thr Leu Asn 2160
Ala Thr 2165	Gly Glu Glu Ile	Ile 2170	Gln Gln Ser Ser	Lys Thr Asp Ala 2175
Ser Ile 2180	Leu Gln Glu Lys	Leu 2185	Gly Ser Leu Asn	Leu Arg Trp Gln 2190
Glu Val 2195	Cys Lys Gln Leu	Ser 2200	Asp Arg Lys Lys	Arg Leu Glu Glu 2205
Gln Lys 2210	Asn Ile Leu Ser	Glu 2215	Phe Gln Arg Asp	Leu Asn Glu Phe 2220
Val Leu 2225	Trp Leu Glu Glu	Ala 2230	Asp Asn Ile Ala	Ser Ile Pro Leu 2235
Glu Pro 2240	Gly Lys Glu Gln	Gln 2245	Leu Lys Glu Lys	Leu Glu Gln Val 2250
Lys Leu 2255	Leu Val Glu Glu	Leu 2260	Pro Leu Arg Gln	Gly Ile Leu Lys 2265
Gln Leu 2270	Asn Glu Thr Gly	Gly 2275	Pro Val Leu Val	Ser Ala Pro Ile 2280
Ser Pro 2285	Glu Glu Gln Asp	Lys 2290	Leu Glu Asn Lys	Leu Lys Gln Thr 2295
Asn Leu 2300	Gln Trp Ile Lys	Val 2305	Ser Arg Ala Leu	Pro Glu Lys Gln 2310
Gly Glu 2315	Ile Glu Ala Gln	Ile 2320	Lys Asp Leu Gly	Gln Leu Glu Lys 2325
Lys Leu 2330	Glu Asp Leu Glu	Glu 2335	Gln Leu Asn His	Leu Leu Leu Trp 2340
Leu Ser 2345	Pro Ile Arg Asn	Gln 2350	Leu Glu Ile Tyr	Asn Gln Pro Asn 2355
Gln Glu 2360	Gly Pro Phe Asp	Val 2365	Gln Glu Thr Glu	Ile Ala Val Gln 2370
Ala Lys	Gln Pro Asp Val	Glu	Glu Ile Leu Ser	Lys Gly Gln His

2375		2380		2385
Leu Tyr Lys Glu Lys Pro Ala Thr Gln Pro Val Lys Arg Lys Leu				
2390		2395		2400
Glu Asp Leu Ser Ser Glu Trp Lys Ala Val Asn Arg Leu Leu Gln				
2405		2410		2415
Glu Leu Arg Ala Lys Gln Pro Asp Leu Ala Pro Gly Leu Thr Thr				
2420		2425		2430
Ile Gly Ala Ser Pro Thr Gln Thr Val Thr Leu Val Thr Gln Pro				
2435		2440		2445
Val Val Thr Lys Glu Thr Ala Ile Ser Lys Leu Glu Met Pro Ser				
2450		2455		2460
Ser Leu Met Leu Glu Val Pro Ala Leu Ala Asp Phe Asn Arg Ala				
2465		2470		2475
Trp Thr Glu Leu Thr Asp Trp Leu Ser Leu Leu Asp Gln Val Ile				
2480		2485		2490
Lys Ser Gln Arg Val Met Val Gly Asp Leu Glu Asp Ile Asn Glu				
2495		2500		2505
Met Ile Ile Lys Gln Lys Ala Thr Met Gln Asp Leu Glu Gln Arg				
2510		2515		2520
Arg Pro Gln Leu Glu Glu Leu Ile Thr Ala Ala Gln Asn Leu Lys				
2525		2530		2535
Asn Lys Thr Ser Asn Gln Glu Ala Arg Thr Ile Ile Thr Asp Arg				
2540		2545		2550
Ile Glu Arg Ile Gln Asn Gln Trp Asp Glu Val Gln Glu His Leu				
2555		2560		2565
Gln Asn Arg Arg Gln Gln Leu Asn Glu Met Leu Lys Asp Ser Thr				
2570		2575		2580
Gln Trp Leu Glu Ala Lys Glu Glu Ala Glu Gln Val Leu Gly Gln				
2585		2590		2595
Ala Arg Ala Lys Leu Glu Ser Trp Lys Glu Gly Pro Tyr Thr Val				
2600		2605		2610
Asp Ala Ile Gln Lys Lys Ile Thr Glu Thr Lys Gln Leu Ala Lys				

2615		2620		2625
Asp Leu Arg Gln Trp Gln Thr Asn Val Asp Val Ala Asn Asp Leu				
2630		2635		2640
Ala Leu Lys Leu Leu Arg Asp Tyr Ser Ala Asp Asp Thr Arg Lys				
2645		2650		2655
Val His Met Ile Thr Glu Asn Ile Asn Ala Ser Trp Arg Ser Ile				
2660		2665		2670
His Lys Arg Val Ser Glu Arg Glu Ala Ala Leu Glu Glu Thr His				
2675		2680		2685
Arg Leu Leu Gln Gln Phe Pro Leu Asp Leu Glu Lys Phe Leu Ala				
2690		2695		2700
Trp Leu Thr Glu Ala Glu Thr Thr Ala Asn Val Leu Gln Asp Ala				
2705		2710		2715
Thr Arg Lys Glu Arg Leu Leu Glu Asp Ser Lys Gly Val Lys Glu				
2720		2725		2730
Leu Met Lys Gln Trp Gln Asp Leu Gln Gly Glu Ile Glu Ala His				
2735		2740		2745
Thr Asp Val Tyr His Asn Leu Asp Glu Asn Ser Gln Lys Ile Leu				
2750		2755		2760
Arg Ser Leu Glu Gly Ser Asp Asp Ala Val Leu Leu Gln Arg Arg				
2765		2770		2775
Leu Asp Asn Met Asn Phe Lys Trp Ser Glu Leu Arg Lys Lys Ser				
2780		2785		2790
Leu Asn Ile Arg Ser His Leu Glu Ala Ser Ser Asp Gln Trp Lys				
2795		2800		2805
Arg Leu His Leu Ser Leu Gln Glu Leu Leu Val Trp Leu Gln Leu				
2810		2815		2820
Lys Asp Asp Glu Leu Ser Arg Gln Ala Pro Ile Gly Gly Asp Phe				
2825		2830		2835
Pro Ala Val Gln Lys Gln Asn Asp Val His Arg Ala Phe Lys Arg				
2840		2845		2850
Glu Leu Lys Thr Lys Glu Pro Val Ile Met Ser Thr Leu Glu Thr				

2855		2860		2865
Val Arg Ile Phe Leu Thr Glu Gln Pro Leu Glu Gly Leu Glu Lys				
2870		2875		2880
Leu Tyr Gln Glu Pro Arg Glu Leu Pro Pro Glu Glu Arg Ala Gln				
2885		2890		2895
Asn Val Thr Arg Leu Leu Arg Lys Gln Ala Glu Glu Val Asn Thr				
2900		2905		2910
Glu Trp Glu Lys Leu Asn Leu His Ser Ala Asp Trp Gln Arg Lys				
2915		2920		2925
Ile Asp Glu Thr Leu Glu Arg Leu Gln Glu Leu Gln Glu Ala Thr				
2930		2935		2940
Asp Glu Leu Asp Leu Lys Leu Arg Gln Ala Glu Val Ile Lys Gly				
2945		2950		2955
Ser Trp Gln Pro Val Gly Asp Leu Leu Ile Asp Ser Leu Gln Asp				
2960		2965		2970
His Leu Glu Lys Val Lys Ala Leu Arg Gly Glu Ile Ala Pro Leu				
2975		2980		2985
Lys Glu Asn Val Ser His Val Asn Asp Leu Ala Arg Gln Leu Thr				
2990		2995		3000
Thr Leu Gly Ile Gln Leu Ser Pro Tyr Asn Leu Ser Thr Leu Glu				
3005		3010		3015
Asp Leu Asn Thr Arg Trp Lys Leu Leu Gln Val Ala Val Glu Asp				
3020		3025		3030
Arg Val Arg Gln Leu His Glu Ala His Arg Asp Phe Gly Pro Ala				
3035		3040		3045
Ser Gln His Phe Leu Ser Thr Ser Val Gln Gly Pro Trp Glu Arg				
3050		3055		3060
Ala Ile Ser Pro Asn Lys Val Pro Tyr Tyr Ile Asn His Glu Thr				
3065		3070		3075
Gln Thr Thr Cys Trp Asp His Pro Lys Met Thr Glu Leu Tyr Gln				
3080		3085		3090
Ser Leu Ala Asp Leu Asn Asn Val Arg Phe Ser Ala Tyr Arg Thr				

3095		3100		3105
Ala Met Lys Leu Arg Arg Leu Gln Lys Ala Leu Cys Leu Asp Leu				
3110		3115		3120
Leu Ser Leu Ser Ala Ala Cys Asp Ala Leu Asp Gln His Asn Leu				
3125		3130		3135
Lys Gln Asn Asp Gln Pro Met Asp Ile Leu Gln Ile Ile Asn Cys				
3140		3145		3150
Leu Thr Thr Ile Tyr Asp Arg Leu Glu Gln Glu His Asn Asn Leu				
3155		3160		3165
Val Asn Val Pro Leu Cys Val Asp Met Cys Leu Asn Trp Leu Leu				
3170		3175		3180
Asn Val Tyr Asp Thr Gly Arg Thr Gly Arg Ile Arg Val Leu Ser				
3185		3190		3195
Phe Lys Thr Gly Ile Ile Ser Leu Cys Lys Ala His Leu Glu Asp				
3200		3205		3210
Lys Tyr Arg Tyr Leu Phe Lys Gln Val Ala Ser Ser Thr Gly Phe				
3215		3220		3225
Cys Asp Gln Arg Arg Leu Gly Leu Leu Leu His Asp Ser Ile Gln				
3230		3235		3240
Ile Pro Arg Gln Leu Gly Glu Val Ala Ser Phe Gly Gly Ser Asn				
3245		3250		3255
Ile Glu Pro Ser Val Arg Ser Cys Phe Gln Phe Ala Asn Asn Lys				
3260		3265		3270
Pro Glu Ile Glu Ala Ala Leu Phe Leu Asp Trp Met Arg Leu Glu				
3275		3280		3285
Pro Gln Ser Met Val Trp Leu Pro Val Leu His Arg Val Ala Ala				
3290		3295		3300
Ala Glu Thr Ala Lys His Gln Ala Lys Cys Asn Ile Cys Lys Glu				
3305		3310		3315
Cys Pro Ile Ile Gly Phe Arg Tyr Arg Ser Leu Lys His Phe Asn				
3320		3325		3330
Tyr Asp Ile Cys Gln Ser Cys Phe Phe Ser Gly Arg Val Ala Lys				

3335		3340		3345
Gly His Lys Met His Tyr	Pro Met Val Glu Tyr	Cys Thr Pro Thr		
3350	3355	3360		
Thr Ser Gly Glu Asp Val	Arg Asp Phe Ala Lys	Val Leu Lys Asn		
3365	3370	3375		
Lys Phe Arg Thr Lys Arg	Tyr Phe Ala Lys His	Pro Arg Met Gly		
3380	3385	3390		
Tyr Leu Pro Val Gln Thr	Val Leu Glu Gly Asp	Asn Met Glu Thr		
3395	3400	3405		
Pro Val Thr Leu Ile Asn	Phe Trp Pro Val Asp	Ser Ala Pro Ala		
3410	3415	3420		
Ser Ser Pro Gln Leu Ser	His Asp Asp Thr His	Ser Arg Ile Glu		
3425	3430	3435		
His Tyr Ala Ser Arg Leu	Ala Glu Met Glu Asn	Ser Asn Gly Ser		
3440	3445	3450		
Tyr Leu Asn Asp Ser Ile	Ser Pro Asn Glu Ser	Ile Asp Asp Glu		
3455	3460	3465		
His Leu Leu Ile Gln His	Tyr Cys Gln Ser Leu	Asn Gln Asp Ser		
3470	3475	3480		
Pro Leu Ser Gln Pro Arg	Ser Pro Ala Gln Ile	Leu Ile Ser Leu		
3485	3490	3495		
Glu Ser Glu Glu Arg Gly	Glu Leu Glu Arg Ile	Leu Ala Asp Leu		
3500	3505	3510		
Glu Glu Glu Asn Arg Asn	Leu Gln Ala Glu Tyr	Asp Arg Leu Lys		
3515	3520	3525		
Gln Gln His Glu His Lys	Gly Leu Ser Pro Leu	Pro Ser Pro Pro		
3530	3535	3540		
Glu Met Met Pro Thr Ser	Pro Gln Ser Pro Arg	Asp Ala Glu Leu		
3545	3550	3555		
Ile Ala Glu Ala Lys Leu	Leu Arg Gln His Lys	Gly Arg Leu Glu		
3560	3565	3570		
Ala Arg Met Gln Ile Leu	Glu Asp His Asn Lys	Gln Leu Glu Ser		

3575		3580		3585
Gln Leu His Arg Leu Arg Gln Leu Leu Glu Gln Pro Gln Ala Glu				
3590		3595		3600
Ala Lys Val Asn Gly Thr Thr Val Ser Ser Pro Ser Thr Ser Leu				
3605		3610		3615
Gln Arg Ser Asp Ser Ser Gln Pro Met Leu Leu Arg Val Val Gly				
3620		3625		3630
Ser Gln Thr Ser Asp Ser Met Gly Glu Glu Asp Leu Leu Ser Pro				
3635		3640		3645
Pro Gln Asp Thr Ser Thr Gly Leu Glu Glu Val Met Glu Gln Leu				
3650		3655		3660
Asn Asn Ser Phe Pro Ser Ser Arg Gly Arg Asn Thr Pro Gly Lys				
3665		3670		3675
Pro Met Arg Glu Asp Thr Met				
3680		3685		

<210> 384

<400> 384
000

<210> 385

<400> 385
000

<210> 386

<400> 386
000

<210> 387

<400> 387
000

<210> 388

<400> 388
000

<210> 389

<400> 389

000

<210> 390

<400> 390
000

<210> 391

<400> 391
000

<210> 392

<400> 392
000

<210> 393

<400> 393
000

<210> 394

<400> 394
000

<210> 395

<400> 395
000

<210> 396

<400> 396
000

<210> 397

<400> 397
000

<210> 398

<400> 398
000

<210> 399

<400> 399
000

<210> 400

<400> 400
000

<210> 401

<400> 401
000

<210> 402

<400> 402
000

<210> 403

<400> 403
000

<210> 404

<400> 404
000

<210> 405

<400> 405
000

<210> 406

<400> 406
000

<210> 407

<400> 407
000

<210> 408

<400> 408
000

<210> 409

<400> 409
000

<210> 410

<400> 410
000

<210> 411

<400> 411
000

<210> 412

<400> 412
000

<210> 413

<400> 413
000

<210> 414

<400> 414
000

<210> 415

<400> 415
000

<210> 416

<400> 416
000

<210> 417

<400> 417
000

<210> 418

<400> 418
000

<210> 419

<400> 419
000

<210> 420

<400> 420
000

<210> 421

<400> 421

000

<210> 422

<400> 422
000

<210> 423

<400> 423
000

<210> 424

<400> 424
000

<210> 425

<400> 425
000

<210> 426

<400> 426
000

<210> 427

<400> 427
000

<210> 428

<400> 428
000

<210> 429

<400> 429
000

<210> 430

<400> 430
000

<210> 431

<400> 431
000

<210> 432

<400> 432
000

<210> 433

<400> 433
000

<210> 434

<400> 434
000

<210> 435

<400> 435
000

<210> 436

<400> 436
000

<210> 437

<400> 437
000

<210> 438

<400> 438
000

<210> 439

<400> 439
000

<210> 440

<400> 440
000

<210> 441

<400> 441
000

<210> 442

<211> 1307

<212> PRT

<213> Acidaminococcus sp.

<400> 442

Met Thr Gln Phe Glu Gly Phe Thr Asn Leu Tyr Gln Val Ser Lys Thr
 1 5 10 15

Leu Arg Phe Glu Leu Ile Pro Gln Gly Lys Thr Leu Lys His Ile Gln
 20 25 30

Glu Gln Gly Phe Ile Glu Glu Asp Lys Ala Arg Asn Asp His Tyr Lys
 35 40 45

Glu Leu Lys Pro Ile Ile Asp Arg Ile Tyr Lys Thr Tyr Ala Asp Gln
 50 55 60

Cys Leu Gln Leu Val Gln Leu Asp Trp Glu Asn Leu Ser Ala Ala Ile
 65 70 75 80

Asp Ser Tyr Arg Lys Glu Lys Thr Glu Glu Thr Arg Asn Ala Leu Ile
 85 90 95

Glu Glu Gln Ala Thr Tyr Arg Asn Ala Ile His Asp Tyr Phe Ile Gly
 100 105 110

Arg Thr Asp Asn Leu Thr Asp Ala Ile Asn Lys Arg His Ala Glu Ile
 115 120 125

Tyr Lys Gly Leu Phe Lys Ala Glu Leu Phe Asn Gly Lys Val Leu Lys
 130 135 140

Gln Leu Gly Thr Val Thr Thr Thr Glu His Glu Asn Ala Leu Leu Arg
 145 150 155 160

Ser Phe Asp Lys Phe Thr Thr Tyr Phe Ser Gly Phe Tyr Glu Asn Arg
 165 170 175

Lys Asn Val Phe Ser Ala Glu Asp Ile Ser Thr Ala Ile Pro His Arg
 180 185 190

Ile Val Gln Asp Asn Phe Pro Lys Phe Lys Glu Asn Cys His Ile Phe
 195 200 205

Thr Arg Leu Ile Thr Ala Val Pro Ser Leu Arg Glu His Phe Glu Asn
 210 215 220

Val Lys Lys Ala Ile Gly Ile Phe Val Ser Thr Ser Ile Glu Glu Val
 225 230 235 240

Phe Ser Phe Pro Phe Tyr Asn Gln Leu Leu Thr Gln Thr Gln Ile Asp
 245 250 255

Leu Tyr Asn Gln Leu Leu Gly Gly Ile Ser Arg Glu Ala Gly Thr Glu
260 265 270

Lys Ile Lys Gly Leu Asn Glu Val Leu Asn Leu Ala Ile Gln Lys Asn
275 280 285

Asp Glu Thr Ala His Ile Ile Ala Ser Leu Pro His Arg Phe Ile Pro
290 295 300

Leu Phe Lys Gln Ile Leu Ser Asp Arg Asn Thr Leu Ser Phe Ile Leu
305 310 315 320

Glu Glu Phe Lys Ser Asp Glu Glu Val Ile Gln Ser Phe Cys Lys Tyr
325 330 335

Lys Thr Leu Leu Arg Asn Glu Asn Val Leu Glu Thr Ala Glu Ala Leu
340 345 350

Phe Asn Glu Leu Asn Ser Ile Asp Leu Thr His Ile Phe Ile Ser His
355 360 365

Lys Lys Leu Glu Thr Ile Ser Ser Ala Leu Cys Asp His Trp Asp Thr
370 375 380

Leu Arg Asn Ala Leu Tyr Glu Arg Arg Ile Ser Glu Leu Thr Gly Lys
385 390 395 400

Ile Thr Lys Ser Ala Lys Glu Lys Val Gln Arg Ser Leu Lys His Glu
405 410 415

Asp Ile Asn Leu Gln Glu Ile Ile Ser Ala Ala Gly Lys Glu Leu Ser
420 425 430

Glu Ala Phe Lys Gln Lys Thr Ser Glu Ile Leu Ser His Ala His Ala
435 440 445

Ala Leu Asp Gln Pro Leu Pro Thr Thr Leu Lys Lys Gln Glu Glu Lys
450 455 460

Glu Ile Leu Lys Ser Gln Leu Asp Ser Leu Leu Gly Leu Tyr His Leu
465 470 475 480

Leu Asp Trp Phe Ala Val Asp Glu Ser Asn Glu Val Asp Pro Glu Phe
485 490 495

Ser Ala Arg Leu Thr Gly Ile Lys Leu Glu Met Glu Pro Ser Leu Ser
500 505 510

Phe Tyr Asn Lys Ala Arg Asn Tyr Ala Thr Lys Lys Pro Tyr Ser Val
515 520 525

Glu Lys Phe Lys Leu Asn Phe Gln Met Pro Thr Leu Ala Ser Gly Trp
530 535 540

Asp Val Asn Lys Glu Lys Asn Asn Gly Ala Ile Leu Phe Val Lys Asn
545 550 555 560

Gly Leu Tyr Tyr Leu Gly Ile Met Pro Lys Gln Lys Gly Arg Tyr Lys
565 570 575

Ala Leu Ser Phe Glu Pro Thr Glu Lys Thr Ser Glu Gly Phe Asp Lys
580 585 590

Met Tyr Tyr Asp Tyr Phe Pro Asp Ala Ala Lys Met Ile Pro Lys Cys
595 600 605

Ser Thr Gln Leu Lys Ala Val Thr Ala His Phe Gln Thr His Thr Thr
610 615 620

Pro Ile Leu Leu Ser Asn Asn Phe Ile Glu Pro Leu Glu Ile Thr Lys
625 630 635 640

Glu Ile Tyr Asp Leu Asn Asn Pro Glu Lys Glu Pro Lys Lys Phe Gln
645 650 655

Thr Ala Tyr Ala Lys Lys Thr Gly Asp Gln Lys Gly Tyr Arg Glu Ala
660 665 670

Leu Cys Lys Trp Ile Asp Phe Thr Arg Asp Phe Leu Ser Lys Tyr Thr
675 680 685

Lys Thr Thr Ser Ile Asp Leu Ser Ser Leu Arg Pro Ser Ser Gln Tyr
690 695 700

Lys Asp Leu Gly Glu Tyr Tyr Ala Glu Leu Asn Pro Leu Leu Tyr His
705 710 715 720

Ile Ser Phe Gln Arg Ile Ala Glu Lys Glu Ile Met Asp Ala Val Glu
725 730 735

Thr Gly Lys Leu Tyr Leu Phe Gln Ile Tyr Asn Lys Asp Phe Ala Lys
740 745 750

Gly His His Gly Lys Pro Asn Leu His Thr Leu Tyr Trp Thr Gly Leu
755 760 765

Phe Ser Pro Glu Asn Leu Ala Lys Thr Ser Ile Lys Leu Asn Gly Gln
 770 775 780

Ala Glu Leu Phe Tyr Arg Pro Lys Ser Arg Met Lys Arg Met Ala His
 785 790 795 800

Arg Leu Gly Glu Lys Met Leu Asn Lys Lys Leu Lys Asp Gln Lys Thr
 805 810 815

Pro Ile Pro Asp Thr Leu Tyr Gln Glu Leu Tyr Asp Tyr Val Asn His
 820 825 830

Arg Leu Ser His Asp Leu Ser Asp Glu Ala Arg Ala Leu Leu Pro Asn
 835 840 845

Val Ile Thr Lys Glu Val Ser His Glu Ile Ile Lys Asp Arg Arg Phe
 850 855 860

Thr Ser Asp Lys Phe Phe Phe His Val Pro Ile Thr Leu Asn Tyr Gln
 865 870 875 880

Ala Ala Asn Ser Pro Ser Lys Phe Asn Gln Arg Val Asn Ala Tyr Leu
 885 890 895

Lys Glu His Pro Glu Thr Pro Ile Ile Gly Ile Asp Arg Gly Glu Arg
 900 905 910

Asn Leu Ile Tyr Ile Thr Val Ile Asp Ser Thr Gly Lys Ile Leu Glu
 915 920 925

Gln Arg Ser Leu Asn Thr Ile Gln Gln Phe Asp Tyr Gln Lys Lys Leu
 930 935 940

Asp Asn Arg Glu Lys Glu Arg Val Ala Ala Arg Gln Ala Trp Ser Val
 945 950 955 960

Val Gly Thr Ile Lys Asp Leu Lys Gln Gly Tyr Leu Ser Gln Val Ile
 965 970 975

His Glu Ile Val Asp Leu Met Ile His Tyr Gln Ala Val Val Val Leu
 980 985 990

Glu Asn Leu Asn Phe Gly Phe Lys Ser Lys Arg Thr Gly Ile Ala Glu
 995 1000 1005

Lys Ala Val Tyr Gln Gln Phe Glu Lys Met Leu Ile Asp Lys Leu
 1010 1015 1020

Asn	Cys	Leu	Val	Leu	Lys	Asp	Tyr	Pro	Ala	Glu	Lys	Val	Gly	Gly
1025						1030					1035			
Val	Leu	Asn	Pro	Tyr	Gln	Leu	Thr	Asp	Gln	Phe	Thr	Ser	Phe	Ala
1040						1045					1050			
Lys	Met	Gly	Thr	Gln	Ser	Gly	Phe	Leu	Phe	Tyr	Val	Pro	Ala	Pro
1055						1060					1065			
Tyr	Thr	Ser	Lys	Ile	Asp	Pro	Leu	Thr	Gly	Phe	Val	Asp	Pro	Phe
1070						1075					1080			
Val	Trp	Lys	Thr	Ile	Lys	Asn	His	Glu	Ser	Arg	Lys	His	Phe	Leu
1085						1090					1095			
Glu	Gly	Phe	Asp	Phe	Leu	His	Tyr	Asp	Val	Lys	Thr	Gly	Asp	Phe
1100						1105					1110			
Ile	Leu	His	Phe	Lys	Met	Asn	Arg	Asn	Leu	Ser	Phe	Gln	Arg	Gly
1115						1120					1125			
Leu	Pro	Gly	Phe	Met	Pro	Ala	Trp	Asp	Ile	Val	Phe	Glu	Lys	Asn
1130						1135					1140			
Glu	Thr	Gln	Phe	Asp	Ala	Lys	Gly	Thr	Pro	Phe	Ile	Ala	Gly	Lys
1145						1150					1155			
Arg	Ile	Val	Pro	Val	Ile	Glu	Asn	His	Arg	Phe	Thr	Gly	Arg	Tyr
1160						1165					1170			
Arg	Asp	Leu	Tyr	Pro	Ala	Asn	Glu	Leu	Ile	Ala	Leu	Leu	Glu	Glu
1175						1180					1185			
Lys	Gly	Ile	Val	Phe	Arg	Asp	Gly	Ser	Asn	Ile	Leu	Pro	Lys	Leu
1190						1195					1200			
Leu	Glu	Asn	Asp	Asp	Ser	His	Ala	Ile	Asp	Thr	Met	Val	Ala	Leu
1205						1210					1215			
Ile	Arg	Ser	Val	Leu	Gln	Met	Arg	Asn	Ser	Asn	Ala	Ala	Thr	Gly
1220						1225					1230			
Glu	Asp	Tyr	Ile	Asn	Ser	Pro	Val	Arg	Asp	Leu	Asn	Gly	Val	Cys
1235						1240					1245			
Phe	Asp	Ser	Arg	Phe	Gln	Asn	Pro	Glu	Trp	Pro	Met	Asp	Ala	Asp
1250						1255					1260			

Ala Asn Gly Ala Tyr His Ile Ala Leu Lys Gly Gln Leu Leu Leu
1265 1270 1275

Asn His Leu Lys Glu Ser Lys Asp Leu Lys Leu Gln Asn Gly Ile
1280 1285 1290

Ser Asn Gln Asp Trp Leu Ala Tyr Ile Gln Glu Leu Arg Asn
1295 1300 1305

<210> 443

<211> 1228

<212> PRT

<213> Lachnospiraceae bacterium ND2006

<400> 443

Ala Ala Ser Lys Leu Glu Lys Phe Thr Asn Cys Tyr Ser Leu Ser Lys
1 5 10 15

Thr Leu Arg Phe Lys Ala Ile Pro Val Gly Lys Thr Gln Glu Asn Ile
20 25 30

Asp Asn Lys Arg Leu Leu Val Glu Asp Glu Lys Arg Ala Glu Asp Tyr
35 40 45

Lys Gly Val Lys Lys Leu Leu Asp Arg Tyr Tyr Leu Ser Phe Ile Asn
50 55 60

Asp Val Leu His Ser Ile Lys Leu Lys Asn Leu Asn Asn Tyr Ile Ser
65 70 75 80

Leu Phe Arg Lys Lys Thr Arg Thr Glu Lys Glu Asn Lys Glu Leu Glu
85 90 95

Asn Leu Glu Ile Asn Leu Arg Lys Glu Ile Ala Lys Ala Phe Lys Gly
100 105 110

Ala Ala Gly Tyr Lys Ser Leu Phe Lys Lys Asp Ile Ile Glu Thr Ile
115 120 125

Leu Pro Glu Ala Ala Asp Asp Lys Asp Glu Ile Ala Leu Val Asn Ser
130 135 140

Phe Asn Gly Phe Thr Thr Ala Phe Thr Gly Phe Phe Asp Asn Arg Glu
145 150 155 160

Asn Met Phe Ser Glu Glu Ala Lys Ser Thr Ser Ile Ala Phe Arg Cys
165 170 175

Ile Asn Glu Asn Leu Thr Arg Tyr Ile Ser Asn Met Asp Ile Phe Glu

180	185	190
Lys Val Asp Ala Ile Phe Asp	Lys His Glu Val Gln Glu Ile Lys Glu	
195	200	205
Lys Ile Leu Asn Ser Asp Tyr Asp Val Glu Asp Phe Phe Glu Gly Glu		
210	215	220
Phe Phe Asn Phe Val Leu Thr Gln Glu Gly Ile Asp Val Tyr Asn Ala		
225	230	235
Ile Ile Gly Gly Phe Val Thr Glu Ser Gly Glu Lys Ile Lys Gly Leu		
245	250	255
Asn Glu Tyr Ile Asn Leu Tyr Asn Ala Lys Thr Lys Gln Ala Leu Pro		
260	265	270
Lys Phe Lys Pro Leu Tyr Lys Gln Val Leu Ser Asp Arg Glu Ser Leu		
275	280	285
Ser Phe Tyr Gly Glu Gly Tyr Thr Ser Asp Glu Glu Val Leu Glu Val		
290	295	300
Phe Arg Asn Thr Leu Asn Lys Asn Ser Glu Ile Phe Ser Ser Ile Lys		
305	310	315
Lys Leu Glu Lys Leu Phe Lys Asn Phe Asp Glu Tyr Ser Ser Ala Gly		
325	330	335
Ile Phe Val Lys Asn Gly Pro Ala Ile Ser Thr Ile Ser Lys Asp Ile		
340	345	350
Phe Gly Glu Trp Asn Leu Ile Arg Asp Lys Trp Asn Ala Glu Tyr Asp		
355	360	365
Asp Ile His Leu Lys Lys Lys Ala Val Val Thr Glu Lys Tyr Glu Asp		
370	375	380
Asp Arg Arg Lys Ser Phe Lys Lys Ile Gly Ser Phe Ser Leu Glu Gln		
385	390	395
Leu Gln Glu Tyr Ala Asp Ala Asp Leu Ser Val Val Glu Lys Leu Lys		
405	410	415
Glu Ile Ile Ile Gln Lys Val Asp Glu Ile Tyr Lys Val Tyr Gly Ser		
420	425	430
Ser Glu Lys Leu Phe Asp Ala Asp Phe Val Leu Glu Lys Ser Leu Lys		

435	440	445										
Lys Asn Asp Ala Val Val Ala Ile Met Lys Asp Leu Leu Asp Ser Val												
450	455	460										
Lys Ser Phe Glu Asn Tyr Ile Lys Ala Phe Phe Gly Glu Gly Lys Glu												
465	470	475										
Thr Asn Arg Asp Glu Ser Phe Tyr Gly Asp Phe Val Leu Ala Tyr Asp												
	485	490										
Ile Leu Leu Lys Val Asp His Ile Tyr Asp Ala Ile Arg Asn Tyr Val												
	500	505										
Thr Gln Lys Pro Tyr Ser Lys Asp Lys Phe Lys Leu Tyr Phe Gln Asn												
	515	520										
Pro Gln Phe Met Gly Gly Trp Asp Lys Asp Lys Glu Thr Asp Tyr Arg												
	530	535										
Ala Thr Ile Leu Arg Tyr Gly Ser Lys Tyr Tyr Leu Ala Ile Met Asp												
545	550	555										
Lys Lys Tyr Ala Lys Cys Leu Gln Lys Ile Asp Lys Asp Asp Val Asn												
	565	570										
Gly Asn Tyr Glu Lys Ile Asn Tyr Lys Leu Leu Pro Gly Pro Asn Lys												
	580	585										
Met Leu Pro Lys Val Phe Phe Ser Lys Lys Trp Met Ala Tyr Tyr Asn												
	595	600										
Pro Ser Glu Asp Ile Gln Lys Ile Tyr Lys Asn Gly Thr Phe Lys Lys												
	610	615										
Gly Asp Met Phe Asn Leu Asn Asp Cys His Lys Leu Ile Asp Phe Phe												
625	630	635										
Lys Asp Ser Ile Ser Arg Tyr Pro Lys Trp Ser Asn Ala Tyr Asp Phe												
	645	650										
Asn Phe Ser Glu Thr Glu Lys Tyr Lys Asp Ile Ala Gly Phe Tyr Arg												
	660	665										
Glu Val Glu Glu Gln Gly Tyr Lys Val Ser Phe Glu Ser Ala Ser Lys												
	675	680										
Lys Glu Val Asp Lys Leu Val Glu Glu Gly Lys Leu Tyr Met Phe Gln												

690		695		700													
Ile	Tyr	Asn	Lys	Asp	Phe	Ser	Asp	Lys	Ser	His	Gly	Thr	Pro	Asn	Leu		
705					710					715					720		
His	Thr	Met	Tyr	Phe	Lys	Leu	Leu	Phe	Asp	Glu	Asn	Asn	His	Gly	Gln		
				725					730					735			
Ile	Arg	Leu	Ser	Gly	Gly	Ala	Glu	Leu	Phe	Met	Arg	Arg	Ala	Ser	Leu		
			740					745					750				
Lys	Lys	Glu	Glu	Leu	Val	Val	His	Pro	Ala	Asn	Ser	Pro	Ile	Ala	Asn		
		755					760					765					
Lys	Asn	Pro	Asp	Asn	Pro	Lys	Lys	Thr	Thr	Thr	Leu	Ser	Tyr	Asp	Val		
	770					775					780						
Tyr	Lys	Asp	Lys	Arg	Phe	Ser	Glu	Asp	Gln	Tyr	Glu	Leu	His	Ile	Pro		
785					790				795						800		
Ile	Ala	Ile	Asn	Lys	Cys	Pro	Lys	Asn	Ile	Phe	Lys	Ile	Asn	Thr	Glu		
			805					810						815			
Val	Arg	Val	Leu	Leu	Lys	His	Asp	Asp	Asn	Pro	Tyr	Val	Ile	Gly	Ile		
			820					825					830				
Asp	Arg	Gly	Glu	Arg	Asn	Leu	Leu	Tyr	Ile	Val	Val	Val	Asp	Gly	Lys		
		835					840					845					
Gly	Asn	Ile	Val	Glu	Gln	Tyr	Ser	Leu	Asn	Glu	Ile	Ile	Asn	Asn	Phe		
	850					855					860						
Asn	Gly	Ile	Arg	Ile	Lys	Thr	Asp	Tyr	His	Ser	Leu	Leu	Asp	Lys	Lys		
865					870					875					880		
Glu	Lys	Glu	Arg	Phe	Glu	Ala	Arg	Gln	Asn	Trp	Thr	Ser	Ile	Glu	Asn		
				885					890					895			
Ile	Lys	Glu	Leu	Lys	Ala	Gly	Tyr	Ile	Ser	Gln	Val	Val	His	Lys	Ile		
			900					905					910				
Cys	Glu	Leu	Val	Glu	Lys	Tyr	Asp	Ala	Val	Ile	Ala	Leu	Glu	Asp	Leu		
		915					920					925					
Asn	Ser	Gly	Phe	Lys	Asn	Ser	Arg	Val	Lys	Val	Glu	Lys	Gln	Val	Tyr		
	930					935					940						
Gln	Lys	Phe	Glu	Lys	Met	Leu	Ile	Asp	Lys	Leu	Asn	Tyr	Met	Val	Asp		

945		950		955		960									
Lys	Lys	Ser	Asn	Pro	Cys	Ala	Thr	Gly	Gly	Ala	Leu	Lys	Gly	Tyr	Gln
			965						970					975	
Ile	Thr	Asn	Lys	Phe	Glu	Ser	Phe	Lys	Ser	Met	Ser	Thr	Gln	Asn	Gly
			980					985					990		
Phe	Ile	Phe	Tyr	Ile	Pro	Ala	Trp	Leu	Thr	Ser	Lys	Ile	Asp	Pro	Ser
		995					1000					1005			
Thr	Gly	Phe	Val	Asn	Leu	Leu	Lys	Thr	Lys	Tyr	Thr	Ser	Ile	Ala	
	1010					1015					1020				
Asp	Ser	Lys	Lys	Phe	Ile	Ser	Ser	Phe	Asp	Arg	Ile	Met	Tyr	Val	
	1025					1030					1035				
Pro	Glu	Glu	Asp	Leu	Phe	Glu	Phe	Ala	Leu	Asp	Tyr	Lys	Asn	Phe	
	1040					1045					1050				
Ser	Arg	Thr	Asp	Ala	Asp	Tyr	Ile	Lys	Lys	Trp	Lys	Leu	Tyr	Ser	
	1055					1060					1065				
Tyr	Gly	Asn	Arg	Ile	Arg	Ile	Phe	Ala	Ala	Ala	Lys	Lys	Asn	Asn	
	1070					1075					1080				
Val	Phe	Ala	Trp	Glu	Glu	Val	Cys	Leu	Thr	Ser	Ala	Tyr	Lys	Glu	
	1085					1090					1095				
Leu	Phe	Asn	Lys	Tyr	Gly	Ile	Asn	Tyr	Gln	Gln	Gly	Asp	Ile	Arg	
	1100					1105					1110				
Ala	Leu	Leu	Cys	Glu	Gln	Ser	Asp	Lys	Ala	Phe	Tyr	Ser	Ser	Phe	
	1115					1120					1125				
Met	Ala	Leu	Met	Ser	Leu	Met	Leu	Gln	Met	Arg	Asn	Ser	Ile	Thr	
	1130					1135					1140				
Gly	Arg	Thr	Asp	Val	Asp	Phe	Leu	Ile	Ser	Pro	Val	Lys	Asn	Ser	
	1145					1150					1155				
Asp	Gly	Ile	Phe	Tyr	Asp	Ser	Arg	Asn	Tyr	Glu	Ala	Gln	Glu	Asn	
	1160					1165					1170				
Ala	Ile	Leu	Pro	Lys	Asn	Ala	Asp	Ala	Asn	Gly	Ala	Tyr	Asn	Ile	
	1175					1180					1185				
Ala	Arg	Lys	Val	Leu	Trp	Ala	Ile	Gly	Gln	Phe	Lys	Lys	Ala	Glu	

1190		1195		1200
Asp Glu Lys Leu Asp Lys Val Lys Ile Ala Ile Ser Asn Lys Glu				
1205		1210		1215
Trp Leu Glu Tyr Ala Gln Thr Ser Val Lys				
1220		1225		

<210> 444

<400> 444
000

<210> 445

<400> 445
000

<210> 446

<400> 446
000

<210> 447

<400> 447
000

<210> 448

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 448

aaaaaggaaa aaagaagaaa aaga

24

<210> 449

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 449

caaaaaggaa aaaagaagaa aaag

24

<210> 450

<211> 24
 <212> DNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 450
 gcaaaaagga aaaaagaaga aaaa 24

<210> 451
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 451
 uuuugcaaaa aggaaaaaag aaga 24

<210> 452
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 452
 uuuuugcaaa aaggaaaaaa gaag 24

<210> 453
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 453
 guuuuugcaa aaaggaaaaa agaa 24

<210> 454
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 454
 auuuuggguu uuugcaaaaa ggaa 24

<210> 455
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 455
 uauuuugggu uuuugcaaaa agga 24

<210> 456
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 456
 auauuuuggg uuuuugcaaa aagg 24

<210> 457
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 457
 aaauuuugg guuuuugcaa aaag 24

<210> 458
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 458
 gcuaaaauau uuuggguuuu ugca 24

<210> 459
 <211> 24
 <212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 459

agcuaaaaua uuuuggguuu uugc

24

<210> 460

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 460

gagcuaaaau auuuuggguu uuug

24

<210> 461

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 461

agaguaacag ucugaguagg agcu

24

<210> 462

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 462

cagaguaaca gucugaguag gagc

24

<210> 463

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 463

gugacacaac cugugguuac uaag 24

<210> 464
 <211> 23
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 464
 gguuacuaag gaaacugcca ucu 23

<210> 465
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 465
 aaggaaacug ccaucuccaa acua 24

<210> 466
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 466
 aucaucaagc agaagguaug agaa 24

<210> 467
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 467
 agcagaaggu augagaaaaa auga 24

<210> 468
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 468
 gcagaaggua ugagaaaaaa ugau 24

 <210> 469
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 469
 uaaaaguugg cagaaguuuu ucuu 24

 <210> 470
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 470
 aaaaguuggc agaaguuuuu cuuu 24

 <210> 471
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 471
 gguggaaaau cuucauuuuu aaga 24

 <210> 472
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 472
 ugguggaaaa ucucauuuu aaag 24

<210> 473
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 473
 uugguggaaa aucuucuuu uaaa 24

<210> 474
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 474
 gugauuggug gaaaaucuc auuu 24

<210> 475
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 475
 cuaggagagu aaagugauug gugg 24

<210> 476
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 476
 ucuaggagag uaaagugauu ggug 24

<210> 477
 <211> 23
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 477
cuggugggaa auggucuagg aga 23

<210> 478
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 478
guagcacacu guuuuauucu uucu 24

<210> 479
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 479
cacacuguuu auuuuuucu caaa 24

<210> 480
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 480
acacuguuu auuuuucuc aaau 24

<210> 481
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 481
cacuguuuau uuuuuucua aaau 24

<210> 482

<211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 482
 augucuuuuu auuugagaaa agau 24

 <210> 483
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 483
 aagccccaug ucuuuuuauu ugag 24

 <210> 484
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 484
 gaagccccau gucuuuuuau uuga 24

 <210> 485
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 485
 guaagauacc aaaaaggcaa aaca 24

 <210> 486
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 486
 uguaagauac caaaaaggca aaac 24

<210> 487
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 487
 cuguaagaua caaaaaaggc aaaa 24

<210> 488
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 488
 guuccuguaa gauaccaaaa aggc 24

<210> 489
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 489
 aguuccugua agauaccaaa aagg 24

<210> 490
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 490
 uccuggaguu ccuguaagau acca 24

<210> 491
 <211> 24
 <212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 491

auccuggagu uccuguaaga uacc

24

<210> 492

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 492

gggaagaaau aaucagcaa uccu

24

<210> 493

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 493

ggaagaaaua auucagcaau ccuc

24

<210> 494

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 494

gaagaaauaa uucagcaauc cuca

24

<210> 495

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 495

aaaacagaug ccaguaauucu acag 24

<210> 496
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 496
 aaacagaugc caguaauucua cagg 24

<210> 497
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 497
 aacagaugcc aguaauucuac agga 24

<210> 498
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 498
 gaaucugcgg uggcaggagg ucug 24

<210> 499
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 499
 aggucugcaa acagcuguca gaca 24

<210> 500
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 500
 ggucugcaaa cagcugucag acag 24

 <210> 501
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 501
 gucugcaaac agcugucaga caga 24

 <210> 502
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 502
 ucugcaaaca gcugucagac agaa 24

 <210> 503
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 503
 uagggcgaca gaucuaauag gaau 24

 <210> 504
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 504
 agggcgacag aucuaauagg aaug 24

<210> 505
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 505
 uaaagaaagc uuaaaaaguc ugcu 24

<210> 506
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 506
 cuaaagaaag cuaaaaagu cugc 24

<210> 507
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 507
 aaauauucuu cuaaagaaag cuua 24

<210> 508
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 508
 gaaauauucu ucuaagaaa gcuu 24

<210> 509
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 509
ugaaaauauuc uucuaaagaa agcu 24

<210> 510
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 510
ucucucauga aaauuucuuc uaaa 24

<210> 511
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 511
auaaucucuc augaaauuu cuuc 24

<210> 512
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 512
gcguauauuu uuugguuaua cuga 24

<210> 513
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 513
ucaagaaaaa uagauggauu augu 24

<210> 514

<211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 514
 aucaagaaaa auagauggau uaug 24

 <210> 515
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 515
 cagguaaaaag cauauggauc aaga 24

 <210> 516
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 516
 gcagguaaaa gcuauggau caag 24

 <210> 517
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 517
 ugcagguaaa agcauaugga ucaa 24

 <210> 518
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 518
caggcgauuu gacagaucug uuga 24

<210> 519
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 519
agaucuguca aaucgccugc aggu 24

<210> 520
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 520
cagaucuguc aaaucgccug cagg 24

<210> 521
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 521
gccgccauuu cucaacagau cugu 24

<210> 522
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 522
aauggcggcg uuuucauuau gaua 24

<210> 523
<211> 24
<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 523

auuaaaauauc uuuaaucau aaug

24

<210> 524

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 524

ugagaauugg gaacaugcua aaau

24

<210> 525

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 525

gguaagucuu ugauuuguuu uuuc

24

<210> 526

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 526

aaauacaauu ucgaaaaaaac aaau

24

<210> 527

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 527

aagauaaaau caauuucgaa aaaa

24

<210> 528

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 528

gcugaagaua aaaucaauuu cgaa

24

<210> 529

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 529

ugcugaagau aaauacaauu ucga

24

<210> 530

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 530

gugcugaaga uaaauacaau uucg

24

<210> 531

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 531

ugugcugaag auaauacaa uuuc

24

<210> 532

<211> 24

<212> RNA

<213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 532
 gcacaucugg acucuuaac uucu 24

 <210> 533
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 533
 uaaagagucc agaugugcug aaga 24

 <210> 534
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 534
 aagaucaggu ucugaagggg gaug 24

 <210> 535
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 535
 uucagaaccu gaucuuaag aagu 24

 <210> 536
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 536
 aaauaauga ugacaacaac aguc 24

<210> 537
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 537
 uaauauaaug augacaacaa cagu 24

<210> 538
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 538
 uuuauuuuuc cuuuuauucu aguu 24

<210> 539
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 539
 aaaggaaaaa uaaauauaua guag 24

<210> 540
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 540
 uuucaacuag aaauaaagga aaaa 24

<210> 541
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 541
auucuuucaa cuagaauaaa agga 24

<210> 542
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 542
aaucuuuca acuagaauaa aagg 24

<210> 543
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 543
gaauucuuuc aacuagaaua aaag 24

<210> 544
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 544
auucugaaau cuuucacua gaau 24

<210> 545
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 545
gauucugaau ucuuucacu agaa 24

<210> 546

<211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 546
 cagaaccgga ggcaacaguu gaau 24

<210> 547
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 547
 ggaggcaaca guugaaugaa augu 24

<210> 548
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 548
 uauacaguag augcaaucca aaag 24

<210> 549
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 549
 gaugcaaucc aaaagaaaau caca 24

<210> 550
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 550
 aaucacagaa accaagguua guau 24

<210> 551
 <211> 23
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 551
 agguuaguau caaagauacc uuu 23

<210> 552
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 552
 gguuagauac aaagauaccu uuuu 24

<210> 553
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 553
 aguaucaaag auaccuuuuu aaaa 24

<210> 554
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 554
 guaucaaaga uaccuuuuua aaau 24

<210> 555
 <211> 24
 <212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 555

uguuuguguc ccaguuugca uuaa

24

<210> 556

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 556

cugggacaca aacauggcaa uuua

24

<210> 557

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 557

acugggacac aaacauggca auuu

24

<210> 558

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 558

aacugggaca caaacauggc aaau

24

<210> 559

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 559

uauuuguuuaa ugcaaacugg gaca 24

<210> 560
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 560
 acaaauaguu ugagaacuau guug 24

<210> 561
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 561
 caaaauaguuu gagaacuaug uugg 24

<210> 562
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 562
 aaauaguuug agaacuau guug 24

<210> 563
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 563
 auaguuugag aacuauguug gaaa 24

<210> 564
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 564
 uaguuugaga acuauguugg aaaa 24

 <210> 565
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 565
 aguuugagaa cuauguugga aaaa 24

 <210> 566
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 566
 uaguucucaa acuauuuguu aaug 24

 <210> 567
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 567
 uauuuuuuuu uccaacauag uucu 24

 <210> 568
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 568
 cuucuucuc caggcuagaa gaac 24

<210> 569
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 569
 cuucuagccu ggagaaagaa gaau 24

<210> 570
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 570
 ucuucuagcc uggagaaaga agaa 24

<210> 571
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 571
 auucuuuugu ucuucuagcc ugga 24

<210> 572
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 572
 caaaagaaua ucuugucaga auuu 24

<210> 573
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 573
cuggaaaaga gcagcaacua aaag 24

<210> 574
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 574
caagucaagg uaauuuuauu uucu 24

<210> 575
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 575
caaauccccc agggccugcu ugca 24

<210> 576
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 576
aggcccuggg ggauuugaga aaau 24

<210> 577
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 577
caggcccugg gggauuugag aaaa 24

<210> 578

<211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 578
 caagcaggcc cugggggauu ugag 24

<210> 579
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 579
 gcaagcaggc ccugggggau uuga 24

<210> 580
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 580
 gcagaaaacc aaugauugaa uuaa 24

<210> 581
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 581
 ggcagaaaac caaugauuga auua 24

<210> 582
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 582
gggcagaaaa ccaaugauug aaau 24

<210> 583
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 583
ugggcagaaa accaaugauu gaau 24

<210> 584
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 584
auuagguuau ucauaguucc uugc 24

<210> 585
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 585
aacuaugaau aaccuaaugg gcag 24

<210> 586
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 586
gaacuaugaa uaaccuaaug ggca 24

<210> 587
<211> 24
<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 587

uauuuccugu uaaauuguuu ucua

24

<210> 588

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 588

gguuuauaga aaacaauuua acag

24

<210> 589

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 589

auacaguaac aucuuuuuuu uuuc

24

<210> 590

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 590

uacaguaaca ucuuuuuuau uuuc

24

<210> 591

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 591

auguuacugu auaaggguuu auag 24

<210> 592
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 592
 gauguuacug uauaaggguu uaua 24

<210> 593
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 593
 cagccaaaac acuuuuagaa auaa 24

<210> 594
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 594
 ccagccaaaa cacuuuuaga aaau 24

<210> 595
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 595
 accagccaaa acacuuuuag aaau 24

<210> 596
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 596
 gaccagccaa aacacuuuuu gaaa 24

 <210> 597
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 597
 gugagaccag ccaaaacacu uuua 24

 <210> 598
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 598
 aauguacuu uacuuuguau uaug 24

 <210> 599
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 599
 auuguacuuu acuuuguauu augu 24

 <210> 600
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 600
 uaaaguacaa uugugagacc agcc 24

<210> 601
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 601
 guaaaguaca auugugagac cagc 24

<210> 602
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 602
 guauuccuuu uacauaauac aaag 24

<210> 603
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 603
 guuguguauu ccuuuuacau aaau 24

<210> 604
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 604
 auccugcauu guugccugua agaa 24

<210> 605
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 605
uuccaacugg ggacgccucu guuc 24

<210> 606
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 606
uuggaagaac ucauuaccgc ugcc 24

<210> 607
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 607
ucauuaccgc ugccccaaaau uga 24

<210> 608
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 608
cucuugauug cuggucuugu uuuu 24

<210> 609
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 609
guuuuuuaac aagcauggga caca 24

<210> 610

<211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 610
 cuuugugugu cccaugcuug uuaa 24

<210> 611
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 611
 gcuuugugug uccaugcuu guua 24

<210> 612
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 612
 ugcuuugugu guccaugcu uguu 24

<210> 613
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 613
 uugcuuugug ugucccaugc uugu 24

<210> 614
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 614
agcaaugc augacaaguu ucaa 24

<210> 615
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 615
gcaaugca ugacaaguuu caau 24

<210> 616
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 616
caaugcau gacaaguuuc aaau 24

<210> 617
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 617
gauaauugaa cuuaaguuuu uauu 24

<210> 618
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 618
auagaaaucc aaauauauau ucac 24

<210> 619
<211> 24
<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 619

auuaagaugu ucaugaauua ucuu

24

<210> 620

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 620

uaaguaaugu guaугсuuuu cugu

24

<210> 621

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 621

aucuucuaac uuccucuuaa acag

24

<210> 622

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 622

gaucuucuaa cuuccucuua aaca

24

<210> 623

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 623

aucugagcuc ugaguggaag gcgg 24

<210> 624
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 624
accguuuacu ucaagagcug aggg 24

<210> 625
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 625
cugcuuugcc cucagcucu gaag 24

<210> 626
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 626
ucucuuggc ucuagcuau uguu 24

<210> 627
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 627
cucuuggc cuagcuauuu guuc 24

<210> 628
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 628
 cacuuuugaa caaauagcua gagc 24

 <210> 629
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 629
 ucacuucaua guugcacuuu uga 24

 <210> 630
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 630
 augaagugau gacuggguga gaga 24

 <210> 631
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 631
 ugaagugaug acugggugag agag 24

 <210> 632
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 632
 aagagaaaaa uauauauaua uaua 24

<210> 633
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 633
 gaauuagcug ucuauagaaa gaga 24

<210> 634
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 634
 ugaauuagcu gucuauagaa agag 24

<210> 635
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 635
 agcuauuca uuuuuuuacu guuu 24

<210> 636
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 636
 augaauuagc ugucuauaga aaga 24

<210> 637
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 637
gcuaauucau uuuuuuacug uuuu 24

<210> 638
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 638
aaaaaaauga auuagcuguc uaua 24

<210> 639
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 639
uuaaaauuuu uauauuacag aaau 24

<210> 640
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 640
uaaaauuuuu auauuacaga auau 24

<210> 641
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 641
auauaaaaau uuuaaaacag uaaa 24

<210> 642

<211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 642
 aaauaauaaaa uuuuuuuuuu guaa 24

<210> 643
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 643
 uaauauaaaa auuuuuuuuu acu 24

<210> 644
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 644
 guaaauauaaa aaauuuuuuuu cagu 24

<210> 645
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 645
 uguaauauaaa aaauuuuuuuu acag 24

<210> 646
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 646
uauauucugu aaauaaaaa uuuu 24

<210> 647
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 647
uuauauucug uaauaaaaa auuu 24

<210> 648
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 648
cagaauauaa aagauagucu acaa 24

<210> 649
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 649
cuauuuuuu uauucuguaa uaua 24

<210> 650
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 650
acuauuuuu auauucugua auau 24

<210> 651
<211> 24
<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 651

gacuaucuuu uauauucugu aaau

24

<210> 652

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 652

cauagcaaga agacagcagc auug

24

<210> 653

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 653

cauuuuguua acuuuuuuccc auug

24

<210> 654

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 654

cauauauuuu ucuugauacu ugca

24

<210> 655

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 655

aaaucauuuc ugcaaguauc aaga 24

<210> 656
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 656
caaaucuuu cugcaaguau caag 24

<210> 657
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 657
acaaucuu ucugcaagua ucaa 24

<210> 658
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 658
auaaaauucua caguucccug aaaa 24

<210> 659
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 659
gaauuuauuu caguaccuc caug 24

<210> 660
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 660
 aauuuauuuc aguaccucc augg 24

<210> 661
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 661
 ugaaauaaau ucuacaguuc ccug 24

<210> 662
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 662
 auuuauuuca guaccucca ugga 24

<210> 663
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 663
 cugaaauaaa uucuacaguu cccu 24

<210> 664
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 664
 uuuauuucag uaccuccau ggaa 24

<210> 665
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 665
 uaccuccau ggaaaaaaga cagg 24

<210> 666
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 666
 accuccaug gaaaaaagac aggg 24

<210> 667
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 667
 ccuccaugg aaaaaagaca ggga 24

<210> 668
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 668
 uuuuuuccau ggaggguacu gaaa 24

<210> 669
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 669
ugucuuuuuu ccauggaggg uacu 24

<210> 670
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 670
ccuugagcaa gaaccaugca aacu 24

<210> 671
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 671
ugcucaagga augcauuuuc uuau 24

<210> 672
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 672
gcucaaggaa ugcauuuucu uaug 24

<210> 673
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 673
ugcauuccuu gagcaagaac caug 24

<210> 674

<211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 674
 gaaaauuuau uuccacaugu aggu 24

<210> 675
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 675
 aaaauuuauu uccacaugua gguc 24

<210> 676
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 676
 aaauuuauuu ccacauguag guca 24

<210> 677
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 677
 cauguggaaa uaaauuuuca uaag 24

<210> 678
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 678
acauguggaa auaaaauuuc auaa 24

<210> 679
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 679
ccacauguag gucaaaaaug uaau 24

<210> 680
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 680
cacauguagg ucaaaaaugu aaug 24

<210> 681
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 681
acauguaggu caaaaaugua auga 24

<210> 682
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 682
acaauuuuga ccuacaugug gaaa 24

<210> 683
<211> 24
<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 683

cauuacauuu uugaccuaca ugug

24

<210> 684

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 684

aaaaauauca uggcuggauu gcaa

24

<210> 685

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 685

gcuggauugc aacaaaccaa cagu

24

<210> 686

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 686

cuggauugca acaaaccaac agug

24

<210> 687

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 687

ccuaugacua uggaugagag cauu 24

<210> 688
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 688
 uagguaagaa gauuacugag acau 24

<210> 689
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 689
 auuacugaga cauuaaaaua cuug 24

<210> 690
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 690
 uuacugagac auuaaaauaac uugu 24

<210> 691
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 691
 ggggaaaaau augucaucag aguc 24

<210> 692
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 692
 caugaucugg aaccuacug ggga 24

 <210> 693
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 693
 acaugaucug gaaccuacu gggg 24

 <210> 694
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 694
 gacaugaucu ggaaccuac uggg 24

 <210> 695
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 695
 uacacacaua cacaagaca aaau 24

 <210> 696
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 696
 uacacauaca cacauacaca aaga 24

<210> 697
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 697
 aacacauaca cauacacaca uaca 24

<210> 698
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 698
 auuccaguca aaaggucug gccu 24

<210> 699
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 699
 uauuccaguc aaauaggucu ggcc 24

<210> 700
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 700
 gcuggcaaac cacacuauc cagu 24

<210> 701
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 701
agucguugug uggcugacug cugg 24

<210> 702
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 702
cgccagauau caauuaggca uaga 24

<210> 703
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 703
aaacuacucg auccugaagg uugg 24

<210> 704
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 704
cauacuaaaa gcagugguag ucca 24

<210> 705
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 705
gaaaacauua aacucuacca uacu 24

<210> 706

<211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 706
 ugaaaacauu aaacucuacc auac 24

<210> 707
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 707
 uuguucauua uccuuuuaga gucu 24

<210> 708
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 708
 aaaggauaau gaacaaaauca aagu 24

<210> 709
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 709
 uauccuuuua gagucucaaa uaua 24

<210> 710
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 710
acucuaaaag gauaaugaac aaau 24

<210> 711
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 711
uuuuagaguc ucaaauauag aaac 24

<210> 712
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 712
uuuagagucu caaaauauaga aacc 24

<210> 713
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 713
uuagagucuc aaauauagaa acca 24

<210> 714
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 714
uugagacucu aaaaggauaa ugaa 24

<210> 715
<211> 24
<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 715

uuugguuucu auauuugaga cucu

24

<210> 716

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 716

uuuugguuuc uauauuugag acuc

24

<210> 717

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 717

agcaugaag ccauccagga agug

24

<210> 718

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 718

gcuucaugc ucacuuguug aggc

24

<210> 719

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 719

ggcuucaaug cucacuuguu gagg 24

<210> 720
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 720
 aguggaaaug uugccaaggc cacc 24

<210> 721
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 721
 guugccaagg ccaccuaaag ugac 24

<210> 722
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 722
 gaagaacauu uucaguuaca ucau 24

<210> 723
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 723
 aucaaaugca cuauucucuaa cagg 24

<210> 724
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 724
 auagugcauu ugaugaugua acug 24

 <210> 725
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 725
 aauagugcau ugaugaugu aacu 24

 <210> 726
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 726
 acuauucuca acagguaaag ugug 24

 <210> 727
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 727
 uaccuaaaaa ugcauauaaa acag 24

 <210> 728
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 728
 auaccuaaaaa augcauauaa aaca 24

<210> 729
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 729
 cacguaauac cuaaaaaugc auau 24

<210> 730
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 730
 gcacguaaau ccuaaaaaaug caua 24

<210> 731
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 731
 auauauaugu gcacguaaau ccua 24

<210> 732
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 732
 uauauauaug ugcacguaau accu 24

<210> 733
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 733
auauauauau gugcacguua uacc 24

<210> 734
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 734
cugcacaaua uuauaguugu ugcu 24

<210> 735
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 735
auaaaaagag aaagauggag gaac 24

<210> 736
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 736
caccuaguga acuccauaaa aaga 24

<210> 737
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 737
auggugcacc uagugaacuc caua 24

<210> 738

<211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 738
 aauggugcac cuagugaacu ccau 24

 <210> 739
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 739
 gaauggugca ccuagugaac ucca 24

 <210> 740
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 740
 gaccaaaugu ucagaugcaa uuau 24

 <210> 741
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

 <400> 741
 ugccucacuc acccugcaaa ggac 24

 <210> 742
 <211> 24
 <212> RNA
 <213> Artificial Sequence

 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 742
agugagcgag aggcugcuuu ggaa 24

<210> 743
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 743
gcagccucuc gcucacucac ccug 24

<210> 744
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 744
uugcaguaau cuaugaguuu cuuc 24

<210> 745
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 745
cugcaacagu uccccugga ccug 24

<210> 746
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 746
ugcaacaguu cccccuggac cugg 24

<210> 747
<211> 24
<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 747

uuucuugccu ggcuuacaga agcu

24

<210> 748

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 748

uuucagcuuc uguaagccag gcaa

24

<210> 749

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 749

guccuacagg augcuacccg uaag

24

<210> 750

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 750

ggcuccuaga agacuccaag ggag

24

<210> 751

<211> 24

<212> RNA

<213> Artificial Sequence

<220>

<221> source

<223> /note="Description of Artificial Sequence: Synthetic oligonucleotide"

<400> 751

gcuccuagaa gacuccaagg gagu 24

<210> 752
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 752
 cuccaaggga guaaaagagc ugau 24

<210> 753
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 753
 uggauccaca agagugcuaa agcg 24

<210> 754
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 754
 guucaauugg auccacaaga gugc 24

<210> 755
 <211> 24
 <212> RNA
 <213> Artificial Sequence
 <220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 755
 uacuuguaac ugacaagcca ggga 24

<210> 756
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 756
 acuuguaacu gacaagccag ggac 24

<210> 757
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 757
 guaacugaca agccagggac aaaa 24

<210> 758
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 758
 uaacugacaa gccagggaca aaac 24

<210> 759
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 759
 ucccuggcuu gucaguuaca agua 24

<210> 760
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 760
 cagaguaaca gucugaguag gagc 24

<210> 761
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 761
 uacuuuguuu agcaauacau ggua 24

<210> 762
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 762
 uggcucaaau uguuacucuu caau 24

<210> 763
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 763
 cuuucaaaga acuuugcaga gccu 24

<210> 764
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source
 <223> /note="Description of Artificial Sequence: Synthetic
 oligonucleotide"

<400> 764
 guugaagcca uuuuguugcu cuuu 24

<210> 765
 <211> 24
 <212> RNA
 <213> Artificial Sequence

<220>
 <221> source

<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 765
guugaagcca uuuuauugcu cuuu 24

<210> 766
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 766
uuuugaggcu cugcaaaguu cuuu 24

<210> 767
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 767
aguuauuaau gcuaagauau ucag 24

<210> 768
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 768
uauaauaugc ccuguaauau aaau 24

<210> 769
<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 769
uaaaggccaa accucggcuu accu 24

<210> 770

<211> 24
<212> RNA
<213> Artificial Sequence

<220>
<221> source
<223> /note="Description of Artificial Sequence: Synthetic
oligonucleotide"

<400> 770
ucaauaucuu ugaaggacuc uggg

24