



(51) International Patent Classification:

C12N 9/22 (2006.01) C12N 15/63 (2006.01)
C12N 15/09 (2006.01) C12N 15/85 (2006.01)

(21) International Application Number:

PCT/US2017/036821

(22) International Filing Date:

09 June 2017 (09.06.2017)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

62/348,262 10 June 2016 (10.06.2016) US

(71) Applicants: CITY OF HOPE [US/US]; 1500 East Duarte Road, Duarte, CA 91010 (US). THE REGENTS OF THE UNIVERSITY OF CALIFORNIA [US/US]; 1111 Franklin Street, Twelfth Floor, Oakland, CA 94607 (US).

(72) Inventors: BURNETT, John; 9626 Lubao Avenue, Chatsworth, CA 91311 (US). PHAM, Anh; 9626 Lubao Avenue, Chatsworth, CA 91311 (US).

(74) Agent: MITRA, Linyu, L. et al.; Mintz Levin Cohn Ferris Glovsky and Popeo, P.C., One Financial Center, Boston, MA 02111 (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).

(54) Title: COMPOSITIONS AND METHODS FOR MITOCHONDRIAL GENOME EDITING

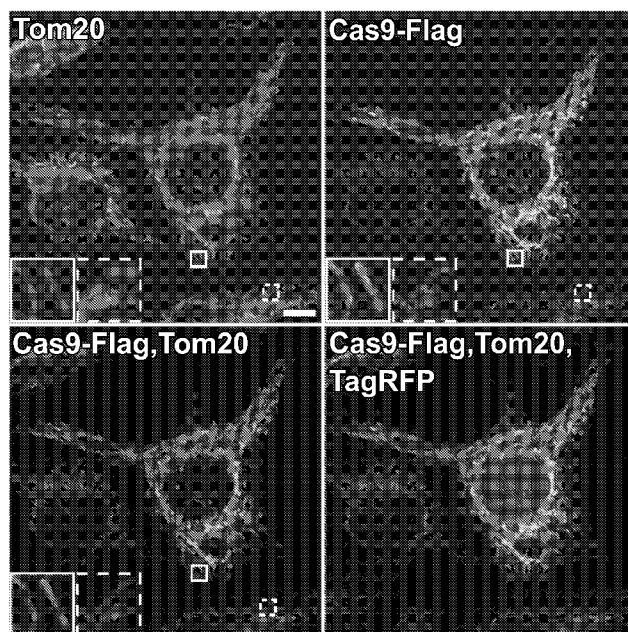


FIG. 1C

(57) Abstract: Compositions and methods for mitochondria genome editing are provided. Also provided are methods for treating mitochondrial disorders by the disclosed compositions.

Published:

- *with international search report (Art. 21(3))*
- *with sequence listing part of description (Rule 5.2(a))*

COMPOSITIONS AND METHODS FOR MITOCHONDRIAL GENOME EDITING

CROSS-REFERENCES TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. Provisional Application No. 62/348,262, filed
5 June 10, 2016, the content of which is incorporated herein by reference in its entirety and for all
purposes.

REFERENCE TO A "SEQUENCE LISTING," A TABLE, OR A COMPUTER PROGRAM LISTING APPENDIX SUBMITTED AS AN ASCII FILE

[0002] The Sequence Listing written in file 48440-598001WO_ST25.TXT, created June 8, 2017,
10 175,122 bytes in size, machine format IBM-PC, MS-Windows operating system, is hereby
incorporated by reference.

BACKGROUND OF THE INVENTION

[0003] Mitochondria are essential organelles that generate the bulk of cellular energy in the form
of ATP from the oxidation of carbohydrates and fats. To carry out this central role in bioenergetics,
15 mitochondria require their own genome, a 16.6 kilobase (kb) circular double-stranded molecule that
encodes 37 genes. Each human cell carries hundreds to thousands of copies of mitochondria DNA
(mtDNA). In mitochondrial encephalomyopathies, cells typically contain a mixture of both
pathogenic and normal mtDNA molecules, a state termed heteroplasmy. There are over 600 known
mtDNA mutations associated with such mtDNA diseases, which have diverse clinical features,
20 including maternal inheritance, defects in the central and peripheral nervous systems, muscle
defects, and exercise intolerance. Due to the inability to transform mitochondrial DNA, there are no
approved clinical therapies for the treatment of mitochondrial diseases. Provided herein are
solutions to these and other problems in the art.

BRIEF SUMMARY OF THE INVENTION

[0004] In a first aspect, there is provided a composition including a delivery vehicle and a protein including a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme, where the protein is bound to the delivery vehicle.

5 **[0005]** In another aspect, there is provided a protein including a mitochondrial localization amino acid sequence covalently attached to Cpf1 or a Class II CRISPR endonuclease or a Cas9 variant.

[0006] In another aspect, there is provided a nucleic acid including a mitochondrial import sequence and a single-guide RNA (sgRNA) sequence.

10 **[0007]** In another aspect, there is provided a method of altering expression of at least one mitochondrial nucleic acid sequence, the method including introducing into an eukaryotic cell a nucleic acid as disclosed herein.

[0008] In another aspect, there is provided a method of treating a mitochondrial disorder in a subject in need thereof, the method including administering to the subject an effective amount of a composition disclosed herein and an effective amount of a nucleic acid disclosed herein.

15 **[0009]** In another aspect, there is provided a method of treating a mitochondrial disorder in a subject in need thereof, the method including administering to the subject an effective amount of a protein disclosed herein or a nucleic acid disclosed herein and an effective amount of another nucleic acid disclosed herein.

20 **[0010]** In another aspect, there is provided a kit including a nucleic acid disclosed herein, and a protein including a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme or a nucleic acid sequence encoding the protein.

BRIEF DESCRIPTION OF THE DRAWINGS

[0011] FIGS. 1A-1D. Mito-Cas9 localizes to mitochondria. (FIG. 1A) Lentiviral vector expression system for mitochondrial localizing sgRNA (mito-sgRNA) and mito-Cas9-P2A-TagRFP. (FIG. 1B) Western blot of mitoCas9-Flag fractionation, “N” nuclear, “M” mitochondria, cytoplasmic “C”. (FIG. 1C) Confocal microscopy of HeLa cells transduced with the vector from (A). MitoCas9-Flag co-localizes with the mitochondrial Tom20 protein, while TagRFP is dispersed throughout the cytoplasm and nucleus. (FIG. 1D) Plasmid map of construct disclosed herein. See Table 8 for

specific sequence information forming FIG. 1D. The plasmid uses the SFFV promoter to express Cox8-Cas9 (mitoCas9)-P2A-TagRFP and the U6 promoter to express the CRISPR RNA with unique mitochondrial localization loops.

[0012] FIGS. 2A-2B. Mitochondrial localization loops. (FIG. 2A) Sequence and structure for five mitoloops. (FIG. 2B) General designs mitoloop-sgRNA chimeras. Target: a guide sequence having sufficient complementarity with a target polynucleotide sequence; repeat/anti-repeat: a tracr mate sequence having sufficient complementarity with tracrRNA sequence; mitoloop(s): mitochondrial import sequence(s).

[0013] FIGS. 3A-3B. Mitoloop-sgRNA chimera exhibits some co-localization with mitochondria. (FIG. 3A) Image of HeLa cell transfected with 488-labeled mito-loop sgRNA and stained with mitochondrial marker Tom20. (FIG. 3B) Magnified images of boxed regions in (FIG. 3A). Microscopy image shows HeLa cells transduced with mtCas9 by lentivirus were subsequently transfected by Lipofectamine with Cy5-labeled F-loop sgRNA. Tom20 is used as mitochondrial marker. Partial co-localization of the F-loop sgRNA with mitochondria is observed in boxed region. Panel of images showed high magnification of boxed region in dual and single channels.

[0014] FIGS. 4A-4C. (FIG. 4A) A schematic of dual promoter lentiviral construct that expresses MTS-Cas9-3XFLAG-P2A-TagRFP from the SFFV promoter and the mito-sgRNA from the U6 promoter. (FIG. 4B) Western blot of nuclear, mitochondrial and cytoplasmic fractions showing enrichment of the mitoCas9-Flag protein in the mitochondrial fraction. Lamin A/C is a nuclear marker, β -tubulin represents the cytosol, and heat shock protein 60 (HSP60) is a mitochondrial protein. (FIG. 4C) Photomicrographs depicting HeLa cells transfected with the mtCas9 construct indicated in FIG. 4A. The MTS is the Cox8 mitochondrial localization signal. Tom20 is an endogenous protein localized in the mitochondrial outer membrane and is stained with an anti-Tom20 fluorescent antibody. Mitochondrial-localizing Cas9 (mtCas9) contains a C-terminal FLAG tag and is stained with a fluorescent anti-FLAG antibody. Scale bar represents 10 μ m.

[0015] FIGS. 5A-5C. These figures illustrate the qPCR of sgRNA. (FIG. 5A) Reverse transcription of mito-sgRNA in RNA extract using RT primer with the addition of a stem loop. (FIG. 5B) The resulting cDNA is amplified using indicated primers (half arrows). (FIG. 5C) An agarose gel of the qPCR products from schematic B from nuclear or mitochondrial fraction of RNA

extracts. Hela (H) are untransduced wild-type controls; L represents cells transduced with LDF mito-sgRNA (expected product 95 bp) and ZDF represent zDF mito-sgRNA (expected product 115 bp).

[0016] FIGS. 6A-6C. These figures illustrate the mitochondrial localization of mitoAsCpf1. (FIG. 6A) Schematic of construct targeting mitoAsCpf1 to mitochondria using COX VIII targeting signal (represented as Mito in this scheme). There is a 3X HA tag at the 3' end of AsCpf1 and TagRFP is a marker of transfection. (FIG. 6B) Plasmid map of the construct in FIG. 6A. (FIG. 6C) Immunofluorescence micrographs of Hela cells transfected with mitoAsCpf1 construct in FIG. 6A. Tom20 is used as a marker of mitochondria and the 3X HA is used to determine localization of mitoAsCpf1. TagRFP expression is not shown. Scale bar represents 10 μ m.

[0017] FIG. 7. Mitochondrial morphology in cytB deletion cybrids with constitutive expression of mitoCas9 and sgRNA mitoloops. Homoplasmic cybrids were transduced with pL_mitoCRISPR vectors (as in FIG. 1D), which express mitoCas9 and the indicated sgRNAs. The sgRNAs include no additional mitoloop (Mock) or sgRNA with LD, RNP, or MRP mitoloops added at the 3' end (as in FIGS. 2A-2B). The transduced cells were sorted by TagRFP, a marker of transduction. The spacer domain is specific for the 14787 cytochrome b deletion mutation. Mitotracker green is a mitochondrial dye that selectively accumulates in healthy mitochondria with intact membrane potential. Cybrids with the sgRNA-MRP loop exhibited depolarized mitochondria as indicated by the diffuse signal of Mitotracker Green. This suggests possible targeted effects on mtDNA thereby resulting in increased susceptibility to transient depolarization of mitochondrial membrane potential. Scale bar represents 10 μ m.

[0018] FIG. 8. Mitochondrial enrichment of sgRNA in Hela cells stably expressing mitoCRISPR with various mitoloops. The control cells have constitutive expression of mitoCas9 and sgRNA without mitoloop additions. The sgRNA RNA levels were measured relative to two housekeeping RNAs, B-actin for the cytoplasm and cytochrome b for the mitochondria and compared between mitoplast fraction and cytoplasmic fraction. Given the multivariable analysis using the $\Delta\Delta C_t$ method, the Euclidean distance from the control was calculated to obtain a single numerical value for fold enrichment as described in the Methods. As seen in the graph, the attachment of the L-loop and D-loop in tandem (LDF) resulted in a significant enrichment of sgRNA in mitoplast relative to cytochrome b and β -actin. These data were consistent with previous observations of mitochondrial

import for smaller, non-CRISPR nuclear-transcribed RNAs, in which the LDF double-loop was much more efficiently localized into mitochondria than either the LD or LF loops. Comparisons were made pairwise relative to control by one tailed *t-test* with Welch correction. ** $p < 0.01$.

[0019] FIGS. 9A-9B. These figures illustrate the measurement of mtDNA copy number in cytochrome b deletion mutants by qPCR. Cybrids have constitutive expression of mtCas9 and various sgRNA-mitoloops. Mitochondrial DNA content was measured by extrapolating the copy numbers of cytochrome b and β -actin genes from standard curves. (FIG. 9A) Graph shows mtDNA per cell normalized to mock control. The sgRNA-ctrl has no mitoloops attached to the sgRNA. (FIG. 9B) Data table of FIG. 9A with the mean and standard deviation listed. The addition of G5S, RNaseP (RNP) and MRP loops show decreased copies of mtDNA relative to mock controls.

[0020] FIGS. 10A-10B. These figures illustrate the measurement of mtDNA copy number using a second set of mitochondrial and nuclear genes. The amount of mtDNA copy per cell is calculated by $\Delta\Delta CT$ method using two regions in the ND1 mitochondrial gene (e.g., ND1-int and ND1-5') and the nuclear HBB gene. (FIG. 10A) Graph shows mtDNA content per nuclei normalized to mock control. The G5S, RNaseP, MRP, and zDF (D and F loops placed in the tetraloop and stemloop 2, respectively) mitoloops show decreased levels of mtDNA relative to mock control. (FIG. 10B) Data table of FIG. 10A showing mean and upper and lower limits of fold change.

Table 10
Primers
ND1-internal (ND1-int) Fwd: CCCTAAAACCCGCCACATCT (SEQ ID NO: 119)
ND1-internal Rev: GAGCGATGGTGAGAGCTAAGGT (SEQ ID NO: 120)
ND1-5' Fwd: AACATACCCATGGCCAACCT (SEQ ID NO: 121)
ND1-5' Rev: AGCGAAGGGTTGTAGTAGCCC (SEQ ID NO: 122)
HBB Fwd: GAAGAGCCAAGGACAGGTAC (SEQ ID NO: 123)
HBB Rev: CAACTTCATCCACGTTACC (SEQ ID NO: 124)

[0021] FIG. 11. Staining of mitochondrial DNA and morphology in cytochrome b deletion cybrids with mtCas9 and various sgRNA-mitoloops. Cells were incubated with Picogreen to stain both nuclear and mitochondrial DNA. Fluorescence micrographs show live cell imaging with DNA labeled by Picogreen (first column) and Mitotracker Red (second column) for mitochondrial morphology. The mtDNA puncta appears fewer in cybrids with the LD, zDF, G5S, RNaseP and MRP sgRNA-mitoloops. The decreased labeling of mtDNA is consistent with the quantitative PCR results. Scale bar is 10 μ m.

DETAILED DESCRIPTION OF THE INVENTION

I. Definitions

[0022] While various embodiments and aspects of the present invention are shown and described herein, it will be obvious to those skilled in the art that such embodiments and aspects are provided by way of example only. Numerous variations, changes, and substitutions will now occur to those

skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention.

[0023] The section headings used herein are for organizational purposes only and are not to be construed as limiting the subject matter described. All documents, or portions of documents, cited in the application including, without limitation, patents, patent applications, articles, books, manuals, and treatises are hereby expressly incorporated by reference in their entirety for any purpose.

[0024] Unless defined otherwise, all technical and scientific terms used herein have the meaning commonly understood by a person skilled in the art to which this invention belongs. The following references provide one of skill with a general definition of many of the terms used in this invention: Singleton et al., Dictionary of Microbiology and Molecular Biology (2nd ed. 1994); The Cambridge Dictionary of Science and Technology (Walker ed., 1988); The Glossary of Genetics, 5th Ed., R. Rieger et al. (eds.), Springer Verlag (1991); and Hale & Marham, The Harper Collins Dictionary of Biology (1991). As used herein, the following terms have the meanings ascribed to them unless specified otherwise.

[0025] The use of a singular indefinite or definite article (e.g., “a,” “an,” “the,” etc.) in this disclosure and in the following claims follows the traditional approach in patents of meaning “at least one” unless in a particular instance it is clear from context that the term is intended in that particular instance to mean specifically one and only one. Likewise, the term “comprising” is open ended, not excluding additional items, features, components, etc. References identified herein are expressly incorporated herein by reference in their entireties unless otherwise indicated.

[0026] The terms “comprise,” “include,” and “have,” and the derivatives thereof, are used herein interchangeably as comprehensive, open-ended terms. For example, use of “comprising,” “including,” or “having” means that whatever element is comprised, had, or included, is not the only element encompassed by the subject of the clause that contains the verb.

[0027] The term “amino acid” refers to naturally occurring and synthetic amino acids, as well as amino acid analogs and amino acid mimetics that function in a manner similar to the naturally occurring amino acids. Naturally occurring amino acids are those encoded by the genetic code, as well as those amino acids that are later modified, *e.g.*, hydroxyproline, γ -carboxyglutamate, and O-

phosphoserine. Amino acid analogs refers to compounds that have the same basic chemical structure as a naturally occurring amino acid, *i.e.*, an α carbon that is bound to a hydrogen, a carboxyl group, an amino group, and an R group, *e.g.*, homoserine, norleucine, methionine sulfoxide, methionine methyl sulfonium. Such analogs have modified R groups (*e.g.*, norleucine) or modified peptide backbones, but retain the same basic chemical structure as a naturally occurring amino acid. Amino acid mimetics refers to chemical compounds that have a structure that is different from the general chemical structure of an amino acid, but that functions in a manner similar to a naturally occurring amino acid. The terms “non-naturally occurring amino acid” and “unnatural amino acid” refer to amino acid analogs, synthetic amino acids, and amino acid mimetics which are not found in nature.

[0028] Amino acids may be referred to herein by either their commonly known three letter symbols or by the one-letter symbols recommended by the IUPAC-IUB Biochemical Nomenclature Commission. Nucleotides, likewise, may be referred to by their commonly accepted single-letter codes.

[0029] The terms “polypeptide,” “peptide” and “protein” are used interchangeably herein to refer to a polymer of amino acid residues, wherein the polymer may in embodiments be conjugated to a moiety that does not consist of amino acids. The terms apply to amino acid polymers in which one or more amino acid residue is an artificial chemical mimetic of a corresponding naturally occurring amino acid, as well as to naturally occurring amino acid polymers and non-naturally occurring amino acid polymers. A “fusion protein” refers to a chimeric protein encoding two or more separate protein sequences that are recombinantly expressed as a single moiety.

[0030] As to amino acid sequences, one of skill will recognize that individual substitutions, deletions or additions to a nucleic acid, peptide, polypeptide, or protein sequence which alters, adds or deletes a single amino acid or a small percentage of amino acids in the encoded sequence is a “conservatively modified variant” where the alteration results in the substitution of an amino acid with a chemically similar amino acid. Conservative substitution tables providing functionally similar amino acids are well known in the art. Such conservatively modified variants are in addition to and do not exclude polymorphic variants, interspecies homologs, and alleles of the invention.

[0031] The following eight groups each contain amino acids that are conservative substitutions for one another:

- 1) Alanine (A), Glycine (G);
 - 2) Aspartic acid (D), Glutamic acid (E);
 - 5 3) Asparagine (N), Glutamine (Q);
 - 4) Arginine (R), Lysine (K);
 - 5) Isoleucine (I), Leucine (L), Methionine (M), Valine (V);
 - 6) Phenylalanine (F), Tyrosine (Y), Tryptophan (W);
 - 7) Serine (S), Threonine (T); and
 - 10 8) Cysteine (C), Methionine (M)
- (see, e.g., Creighton, *Proteins* (1984)).

[0032] As may be used herein, the terms “nucleic acid,” “nucleic acid molecule,” “nucleic acid oligomer,” “oligonucleotide,” “nucleic acid sequence,” “nucleic acid fragment” and “polynucleotide” are used interchangeably and are intended to include, but are not limited to, a
15 polymeric form of nucleotides covalently linked together that may have various lengths, either deoxyribonucleotides or ribonucleotides, or analogs, derivatives or modifications thereof. Different polynucleotides may have different three-dimensional structures, and may perform various functions, known or unknown. Non-limiting examples of polynucleotides include a gene, a gene fragment, an exon, an intron, intergenic DNA (including, without limitation, heterochromatic DNA),
20 messenger RNA (mRNA), transfer RNA, ribosomal RNA, a ribozyme, cDNA, a recombinant polynucleotide, a branched polynucleotide, a plasmid, a vector, isolated DNA of a sequence, isolated RNA of a sequence, a nucleic acid probe, and a primer. Polynucleotides useful in the methods of the invention may comprise natural nucleic acid sequences and variants thereof, artificial nucleic acid sequences, or a combination of such sequences.

25 [0033] A polynucleotide is typically composed of a specific sequence of four nucleotide bases: adenine (A); cytosine (C); guanine (G); and thymine (T) (uracil (U) for thymine (T) when the polynucleotide is RNA). Thus, the term “polynucleotide sequence” is the alphabetical representation of a polynucleotide molecule; alternatively, the term may be applied to the polynucleotide molecule itself. This alphabetical representation can be input into databases in a computer having a central
30 processing unit and used for bioinformatics applications such as functional genomics and homology

searching. Polynucleotides may optionally include one or more non-standard nucleotide(s), nucleotide analog(s) and/or modified nucleotides.

[0034] “Nucleic acid” refers to deoxyribonucleotides or ribonucleotides and polymers thereof in either single-, double- or multiple-stranded form, or complements thereof. The term

5 “polynucleotide” refers to a linear sequence of nucleotides. The term “nucleotide” typically refers to a single unit of a polynucleotide, *i.e.*, a monomer. Nucleotides can be ribonucleotides, deoxyribonucleotides, or modified versions thereof. Examples of polynucleotides contemplated herein include single and double stranded DNA, single and double stranded RNA (including siRNA), and hybrid molecules having mixtures of single and double stranded DNA and RNA.

10 Nucleic acids can be linear or branched. For example, nucleic acids can be a linear chain of nucleotides or the nucleic acids can be branched, e.g., such that the nucleic acids comprise one or more arms or branches of nucleotides. Optionally, the branched nucleic acids are repetitively branched to form higher ordered structures such as dendrimers and the like.

[0035] Nucleic acids, including nucleic acids with a phosphothioate backbone can include one or
15 more reactive moieties. As used herein, the term reactive moiety includes any group capable of reacting with another molecule, e.g., a nucleic acid or polypeptide through covalent, non-covalent or other interactions. By way of example, the nucleic acid can include an amino acid reactive moiety that reacts with an amino acid on a protein or polypeptide through a covalent, non-covalent or other interaction.

20 **[0036]** The terms also encompass nucleic acids containing known nucleotide analogues or modified backbone residues or linkages, which are synthetic, naturally occurring, and non-naturally occurring, which have similar binding properties as the reference nucleic acid, and which are metabolized in a manner similar to the reference nucleotides. In other words, the term “nucleotide analog” as used herein generally refers to a purine or pyrimidine nucleotide that differs structurally
25 from A, T, G, C, or U, but is sufficiently similar to substitute for such “normal” nucleotides in a nucleic acid molecule. As used herein, the term “nucleotide analog” encompasses altered bases, different (or unusual) sugars, altered phosphate backbones, or any combination of these alterations. Examples of such analogues include, without limitation, phosphodiester derivatives including, *e.g.*, phosphoramidate, phosphorodiamidate, phosphorothioate (also known as phosphothioate),
30 phosphorodithioate, phosphonocarboxylic acids, phosphonocarboxylates, phosphonoacetic acid,

phosphonoformic acid, methyl phosphonate, boron phosphonate, or O-methylphosphoroamidite linkages (see Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press); and peptide nucleic acid backbones and linkages. Other analogue nucleic acids include those with positive backbones; non-ionic backbones, modified sugars, and non-ribose backbones (e.g. 5 phosphorodiamidate morpholino oligos or locked nucleic acids (LNA)), including those described in U.S. Patent Nos. 5,235,033 and 5,034,506, and Chapters 6 and 7, ASC Symposium Series 580, *Carbohydrate Modifications in Antisense Research*, Sanghui & Cook, eds. Nucleic acids containing one or more carbocyclic sugars are also included within one definition of nucleic acids.

Modifications of the ribose-phosphate backbone may be done for a variety of reasons, e.g., to 10 increase the stability and half-life of such molecules in physiological environments or as probes on a biochip. Nucleotide analogues used herein also include nucleotides having modified 2' position of the ribose ring. For example, the 2' position of the ribose ring is substituted by O-methyl, O-alkyl, O-allyl, S-alkyl, S-allyl, or halo group. Mixtures of naturally occurring nucleic acids and analogues can be made; alternatively, mixtures of different nucleic acid analogues, and mixtures of naturally 15 occurring nucleic acids and analogues may be made. In embodiments, the internucleotide linkages in DNA are phosphodiester, phosphodiester derivatives, or a combination of both.

[0037] “Conservatively modified variants” applies to both amino acid and nucleic acid sequences. With respect to particular nucleic acid sequences, “conservatively modified variants” refers to those nucleic acids that encode identical or essentially identical amino acid sequences. Because of the 20 degeneracy of the genetic code, a number of nucleic acid sequences will encode any given protein. For instance, the codons GCA, GCC, GCG and GCU all encode the amino acid alanine. Thus, at every position where an alanine is specified by a codon, the codon can be altered to any of the corresponding codons described without altering the encoded polypeptide. Such nucleic acid variations are “silent variations,” which are one species of conservatively modified variations.

25 Every nucleic acid sequence herein which encodes a polypeptide also describes every possible silent variation of the nucleic acid. One of skill will recognize that each codon in a nucleic acid (except AUG, which is ordinarily the only codon for methionine, and TGG, which is ordinarily the only codon for tryptophan) can be modified to yield a functionally identical molecule. Accordingly, each silent variation of a nucleic acid which encodes a polypeptide is implicit in each described sequence.

[0038] As used herein, the term “conjugate” refers to the association between atoms or molecules. The association can be direct or indirect. For example, a conjugate between a first moiety (e.g., nucleic acid moiety) and a second moiety (peptide moiety) provided herein can be direct, e.g., by covalent bond, or indirect, e.g., by non-covalent bond (e.g. electrostatic interactions (e.g. ionic bond, hydrogen bond, halogen bond), van der Waals interactions (e.g. dipole-dipole, dipole-induced dipole, London dispersion), ring stacking (pi effects), hydrophobic interactions and the like). In embodiments, conjugates are formed using conjugate chemistry including, but are not limited to nucleophilic substitutions (e.g., reactions of amines and alcohols with acyl halides, active esters), electrophilic substitutions (e.g., enamine reactions) and additions to carbon-carbon and carbon-heteroatom multiple bonds (e.g., Michael reaction, Diels-Alder addition). These and other useful reactions are discussed in, for example, March, ADVANCED ORGANIC CHEMISTRY, 3rd Ed., John Wiley & Sons, New York, 1985; Hermanson, BIOCONJUGATE TECHNIQUES, Academic Press, San Diego, 1996; and Feeney et al., MODIFICATION OF PROTEINS; Advances in Chemistry Series, Vol. 198, American Chemical Society, Washington, D.C., 1982. In embodiments, the first moiety (e.g., polyamine moiety) is non-covalently attached to the second moiety (peptide moiety) through a non-covalent chemical reaction between a component of the first moiety (e.g., polyamine moiety) and a component of the second moiety (peptide moiety). In other embodiments, the first moiety (e.g., polyamine moiety) includes one or more reactive moieties, e.g., a covalent reactive moiety, as described herein (e.g., alkyne, azide, maleimide or thiol reactive moiety). In other embodiments, the first moiety (e.g., polyamine moiety) includes a linker with one or more reactive moieties, e.g., a covalent reactive moiety, as described herein (e.g., alkyne, azide, maleimide or thiol reactive moiety). In other embodiments, the second moiety (peptide moiety) includes one or more reactive moieties, e.g., a covalent reactive moiety, as described herein (e.g., alkyne, azide, maleimide or thiol reactive moiety). In other embodiments, the second moiety (peptide moiety) includes a linker with one or more reactive moieties, e.g., a covalent reactive moiety, as described herein (e.g., alkyne, azide, maleimide or thiol reactive moiety).

[0039] A “labeled nucleic acid or oligonucleotide” is one that is bound, either covalently, through a linker or a chemical bond, or noncovalently, through ionic, van der Waals, electrostatic, or hydrogen bonds to a label such that the presence of the nucleic acid may be detected by detecting the presence of the detectable label bound to the nucleic acid. Alternatively, a method using high affinity

interactions may achieve the same results where one of a pair of binding partners binds to the other, e.g., biotin, streptavidin. In embodiments, the phosphorothioate nucleic acid or phosphorothioate polymer backbone includes a detectable label, as disclosed herein and generally known in the art.

[0040] A “label” or a “detectable moiety” is a composition detectable by spectroscopic, photochemical, biochemical, immunochemical, chemical, or other physical means. For example, useful labels include ³²P, fluorescent dyes, electron-dense reagents, enzymes (e.g., as commonly used in an ELISA), biotin, digoxigenin, or haptens and proteins or other entities which can be made detectable, e.g., by incorporating a radiolabel into a peptide or antibody specifically reactive with a target peptide. Any appropriate method known in the art for conjugating an antibody to the label may be employed, e.g., using methods described in Hermanson, *Bioconjugate Techniques* 1996, Academic Press, Inc., San Diego.

[0041] The term “probe” or “primer”, as used herein, is defined to be one or more nucleic acid fragments whose specific hybridization to a sample can be detected. A probe or primer can be of any length depending on the particular technique it will be used for. For example, PCR primers are generally between 10 and 40 nucleotides in length, while nucleic acid probes for, e.g., a Southern blot, can be more than a hundred nucleotides in length. The probe may be unlabeled or labeled as described below so that its binding to the target or sample can be detected. The probe can be produced from a source of nucleic acids from one or more particular (preselected) portions of a chromosome, e.g., one or more clones, an isolated whole chromosome or chromosome fragment, or a collection of polymerase chain reaction (PCR) amplification products. The length and complexity of the nucleic acid fixed onto the target element is not critical to the invention. One of skill can adjust these factors to provide optimum hybridization and signal production for a given hybridization procedure, and to provide the required resolution among different genes or genomic locations.

[0042] The probe may also be isolated nucleic acids immobilized on a solid surface (e.g., nitrocellulose, glass, quartz, fused silica slides), as in an array. In some embodiments, the probe may be a member of an array of nucleic acids as described, for instance, in WO 96/17958. Techniques capable of producing high density arrays can also be used for this purpose (see, e.g., Fodor (1991) *Science* 767-773; Johnston (1998) *Curr. Biol.* 8: R171-R174; Schummer (1997) *Biotechniques* 23: 1087-1092; Kern (1997) *Biotechniques* 23: 120-124; U.S. Patent No. 5,143,854).

[0043] The term “isolated”, when applied to a nucleic acid or protein, denotes that the nucleic acid or protein is essentially free of other cellular components with which it is associated in the natural state. It can be, for example, in a homogeneous state and may be in either a dry or aqueous solution. Purity and homogeneity are typically determined using analytical chemistry techniques such as polyacrylamide gel electrophoresis or high performance liquid chromatography. A protein that is the predominant species present in a preparation is substantially purified.

[0044] The term “complementary” or “complementarity” refers to the ability of a nucleic acid to form hydrogen bond(s) with another nucleic acid sequence by either traditional Watson-Crick or other non-traditional types. For example, the sequence A-G-T is complementary to the sequence T-C-A. A percent complementarity indicates the percentage of residues in a nucleic acid molecule which can form hydrogen bonds (e.g., Watson-Crick base pairing) with a second nucleic acid sequence (e.g., 5, 6, 7, 8, 9, 10 out of 10 being 50%, 60%, 70%, 80%, 90%, and 100% complementary). “Perfectly complementary” means that all the contiguous residues of a nucleic acid sequence will hydrogen bond with the same number of contiguous residues in a second nucleic acid sequence. “Substantially complementary” as used herein refers to a degree of complementarity that is at least 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99%, or 100% over a region of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 30, 35, 40, 45, 50, or more nucleotides, or refers to two nucleic acids that hybridize under stringent conditions.

[0045] As used herein, “stringent conditions” for hybridization refer to conditions under which a nucleic acid having complementarity to a target sequence predominantly hybridizes with the target sequence, and substantially does not hybridize to non-target sequences. Stringent conditions are generally sequence-dependent, and vary depending on a number of factors. In general, the longer the sequence, the higher the temperature at which the sequence specifically hybridizes to its target sequence. Non-limiting examples of stringent conditions are described in detail in Tijssen (1993), Laboratory Techniques In Biochemistry And Molecular Biology-Hybridization With Nucleic Acid Probes Part 1, Second Chapter “Overview of principles of hybridization and the strategy of nucleic acid probe assay”, Elsevier, N.Y.

[0046] “Hybridization” refers to a reaction in which one or more polynucleotides react to form a complex that is stabilized via hydrogen bonding between the bases of the nucleotide residues. The hydrogen bonding may occur by Watson Crick base pairing, Hoogstein binding, or in any other

sequence specific manner. The complex may comprise two strands forming a duplex structure, three or more strands forming a multi stranded complex, a single self 17 hybridizing strand, or any combination of these. A hybridization reaction may constitute a step in a more extensive process, such as the initiation of PCR, or the cleavage of a polynucleotide by an enzyme. A sequence capable of hybridizing with a given sequence is referred to as the “complement” of the given sequence.

[0047] “Percentage of sequence identity” is determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid base or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity.

[0048] The terms “identical” or percent “identity,” in the context of two or more nucleic acids or polypeptide sequences, refer to two or more sequences or subsequences that are the same or have a specified percentage of amino acid residues or nucleotides that are the same (*i.e.*, 60% identity, optionally 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98%, or 99% identity over a specified region, e.g., of the entire polypeptide sequences of the invention or individual domains of the polypeptides of the invention), when compared and aligned for maximum correspondence over a comparison window, or designated region as measured using one of the following sequence comparison algorithms or by manual alignment and visual inspection. Such sequences are then said to be “substantially identical.” This definition also refers to the complement of a test sequence. Optionally, the identity exists over a region that is at least about 50 nucleotides in length, or more preferably over a region that is 100 to 500 or 1000 or more nucleotides in length.

[0049] An amino acid or nucleotide base “position” is denoted by a number that sequentially identifies each amino acid (or nucleotide base) in the reference sequence based on its position relative to the N-terminus (or 5'-end). Due to deletions, insertions, truncations, fusions, and the like that must be taken into account when determining an optimal alignment, in general the amino acid residue number in a test sequence determined by simply counting from the N-terminus will not

necessarily be the same as the number of its corresponding position in the reference sequence. For example, in a case where a variant has a deletion relative to an aligned reference sequence, there will be no amino acid in the variant that corresponds to a position in the reference sequence at the site of deletion. Where there is an insertion in an aligned reference sequence, that insertion will not
5 correspond to a numbered amino acid position in the reference sequence. In the case of truncations or fusions there can be stretches of amino acids in either the reference or aligned sequence that do not correspond to any amino acid in the corresponding sequence.

[0050] The terms “numbered with reference to” or “corresponding to,” when used in the context of the numbering of a given amino acid or polynucleotide sequence, refers to the numbering of the
10 residues of a specified reference sequence when the given amino acid or polynucleotide sequence is compared to the reference sequence.

[0051] The term “gene” means the segment of DNA involved in producing a protein; it includes regions preceding and following the coding region (leader and trailer) as well as intervening sequences (introns) between individual coding segments (exons). The leader, the trailer as well as
15 the introns include regulatory elements that are necessary during the transcription and the translation of a gene. Further, a “protein gene product” is a protein expressed from a particular gene.

[0052] The word “expression” or “expressed” as used herein in reference to a gene means the transcriptional and/or translational product of that gene. The level of expression of a DNA molecule in a cell may be determined on the basis of either the amount of corresponding mRNA that is present
20 within the cell or the amount of protein encoded by that DNA produced by the cell. The level of expression of non-coding nucleic acid molecules (e.g., sgRNA) may be detected by standard PCR or Northern blot methods well known in the art. *See, Sambrook et al., 1989 Molecular Cloning: A Laboratory Manual*, 18.1-18.88. In illustrative embodiments, sgRNA can be detected by a specificized stem-loop primers. The multiple stem loop structures in SpCas9 RNA provide
25 significant thermostability during reverse transcription (RT) and PCR. To detect the mito-loop sgRNA targets by qPCR, the qPCR method for amplifying miRNA towards sgRNA was adapted. The concept is illustrated in FIGS. 5A-5C. A 44 nucleotide stem loop is attached to the 5' end of the RT primer (FIG. 5A). After reverse transcription, the cDNA is amplified using primers that overlap the target/spacer domain and the stem loop of the RT primer, denoted by the half arrows. A Taqman
30 PCR probe overlaps both portions of the sgRNA and the stem loop (FIG. 5B). FIG. 5C shows the

amplified product of the cDNA after qPCR from a nuclear and mitochondrial fractionation of samples that are untransduced with mitoloops (H) or transduced with LDF (labeled L) or zDF (labeled Z) mito-sgRNA. Exemplary primers are listed in the table below.

RT primer	5'-GTCGTATCCAGTGC GAATACCTCGGACCCTGCACTGGA TACGACCGGACTAGCCTT-3' (SEQ ID NO: 109)
qPCR_R	5'-TACCTCGGACCCTGCACTGG-3' (SEQ ID NO: 110)
qPCR_14787_F	5'-GGTTAGTTTTATTAGGGTTTTAGAGC-3' (SEQ ID NO: 111)
TaqMan probe	5'-AGCAAGTTAAAATAAGGCTAGTCCGGTCGT-3' (SEQ ID NO: 112)

Sequence in bold represents addition of stemloop.

5 **[0053]** The term “mitochondrial localization sequence” or “mitochondria targeting signal” and the like refer, in the usual and customary sense, to a short peptide sequence (about 3-70 amino acids long) that directs a newly synthesized proteins to the mitochondria within a cell. It is usually found at the N-terminus and consists of an alternating pattern of hydrophobic and positively charged amino acids to form what is called an amphipathic helix. Mitochondrial localization sequences can contain
10 additional signals that subsequently target the protein to different regions of the mitochondria, such as the mitochondrial matrix. One exemplary mitochondrial localization sequence is the mitochondrial localization sequence derived from Cox8.

[0054] The term “Cox8” and the like refer, in the usual and customary sense, to cytochrome c oxidase subunit VIII. As used herein, the term refers to both biomolecules having the sequence of
15 Cox8 and truncated and substituted versions thereof, including proteins and nucleic acids encoding the proteins, and to a sequence that performs the function of the cytochrome c oxidase subunit VIII, which function, as known in the art, is coupling of the transfer of electrons from cytochrome c to molecule oxygen, as the terminal step of the respiratory chain. The sequence of Cox8 (GenBank J04823.1) follows:

20 GGCTACGGCTGACCGTTTTTGTGGTGTACTCCGTGCCATCATGTCCGTCTGACGCCGCTGCTGCTGCGGGGCTTGACA
GGCTCGGCCCCGGCGGCTCCAGTGCCGCGCGCCAAGATCCATTTCGTTGCCGCCGAGGGGAAGCTTGGGATCATGGAATT
GGCCGTTGGGCTTACCTCCTGCTTCGTGACCTTCCTCCTGCCAGCGGGCTGGATCCTGTACACCTGGAGACCTACAGGA
GGCCAGAGTGAAGGGGTCCGTTCTGTCCCTCACACTGTGACCTGACCAGCCCCACCGGCCCATCCTGGTCATGTTACTGC
ATTTGTGGCCGGCCTCCCCTGGATCATGTCATTCAATTCCAGTCACCTCTTCTGCAATCATGACCTCTTGATGTCTCCAT
25 GGTGACCTCCTTGGGGGTCACTGACCCTGCTTGGTGGGGTCCCCCTTGTAACAATAAATCTATTTAACTTT (SEQ ID
NO: 54) [Underlined sequence: the mitochondrial localization sequence of Cox8]

[0055] In embodiments, a mitochondrial localization sequence derived from Cox8 includes the amino acid sequence: MSVLTPLLLRGLTGSARRLPVPRAKIHSL (SEQ ID NO: 1). In the embodiments, the mitochondrial localization sequence derived from Cox8 includes an amino acid sequence that is about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% or 95% identity to SEQ ID NO:1.

[0056] In embodiments, a mitochondrial localization sequence derived from Cox8 includes the nucleic acid sequence:
 ATGTCCGTCCTGACGCCGCTGCTGCTGCGGGGCTTGACAGGCTCGGCCCCGGCGGCTCCC
 AGTGCCGCGCGCCAAGATCCATTCGTTG (SEQ ID NO: 2). In embodiments, the mitochondrial localization sequence derived from Cox8 includes a nucleic acid sequence that is about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% or 95% identity to SEQ ID NO:2.

[0057] A “nuclear localization sequence” or “nuclear localization signal (NLS)” is a peptide that directs proteins to the nucleus. In embodiments, the NLS includes five basic, positively charged amino acids. The NLS may be located anywhere on the peptide chain.

[0058] A “mitochondrial import sequence” or “mitochondrial RNA import sequence” refers to an RNA sequence (e.g. small RNAs) capable of directing importation of exogenous RNA to mitochondria. The mitochondrial import sequence may be isolated from or derived from the mitochondrial transcriptome. Exemplary mitochondrial import sequences include, but are not limited to, the 5S ribosomal RNA, RNaseP (“RNP”) and MRP RNA components and a modified γ domain of 5S rRNA, including D loop, F loop, MRP loop, RNP loop, γ 5s loop including the sequence indicated in Table 1 below (see also FIG. 2A). Detailed description of small RNAs as mitochondrial import sequences can be found in Schneider et al., *Annu Rev Biochem.*, 2011, 80:1033-53; Wang et al., *Cell*, 2010, 142(3):456-67; Comte et al., *Nucleic Acids Res*, 2013, 41(1):418-33; Tonin et al., *J Biol Chem*, 2014, 289(19): 13323-34; Smironov et al., *RNA*, 2008, 14(4):749-59; and Zelenka et al., *J Bioenerg Biomembr*, 2014, 46(2):147-56, contents of each of which are incorporated herein as entireties.

Table 1	
mitochondrial import sequence	Sequence
D loop	GCGCAATCGGTAGCGC (SEQ ID NO: 3)
F loop	GAGCCCCCTACAGGGCTC (SEQ ID NO: 4)
MRP	AGAAGCGTATCCCGCTGAGC (SEQ ID NO: 5)
RNP loop	TCTCCCTGAGCTTCAGGGAG (SEQ ID NO: 6)
γ 5s loop	GGCCTGGTTAGTACTTGGATGGGAGACCGCCAAGGAATACCGGGTG (SEQ ID NO: 7)

[0059] In embodiments, the mitochondrial import sequence includes a nucleic acid sequence that is about 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 3, 4, 5, 6 or 7.

5 **[0060]** For specific proteins described herein (e.g., Cas9, Cpf1, and the like), the named protein includes any of the protein's naturally occurring forms, or variants or homologs that maintain the protein transcription factor activity (e.g., within at least 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or 100% activity compared to the native protein). In some embodiments, variants or homologs have at least 90%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity across the
10 whole sequence or a portion of the sequence (e.g. a 50, 100, 150 or 200 continuous amino acid portion) compared to a naturally occurring form. In other embodiments, the protein is the protein as identified by its NCBI sequence reference. In other embodiments, the protein is the protein as identified by its NCBI sequence reference or functional fragment or homolog thereof.

[0061] Thus, a "CRISPR associated protein 9," "Cas9," "Csn1" or "Cas9 protein" as referred to
15 herein includes any of the recombinant or naturally-occurring forms of the Cas9 endonuclease or variants or homologs thereof that maintain Cas9 endonuclease enzyme activity (e.g. within at least 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or 100% activity compared to Cas9). In some aspects, the variants or homologs have at least 90%, 95%, 96%, 97%, 98%, 99% or 100% amino acid
20 sequence identity across the whole sequence or a portion of the sequence (e.g. a 50, 100, 150 or 200 continuous amino acid portion) compared to a naturally occurring Cas9 protein. In embodiments, the Cas9 protein is substantially identical to the protein identified by the UniProt reference number Q99ZW2 or a variant or homolog having substantial identity thereto. Cas9 refers to the protein also known in the art as "nickase". In embodiments, Cas9 is an RNA-guided DNA endonuclease enzyme that binds a CRISPR (clustered regularly interspaced short palindromic repeats) nucleic acid

sequence. In embodiments, the CRISPR nucleic acid sequence is a prokaryotic nucleic acid sequence. In embodiments, the Cas9 nuclease from *Streptococcus pyogenes* is targeted to genomic DNA by a synthetic guide RNA consisting of a 20-nt guide sequence and a scaffold. The guide sequence base-pairs with the DNA target, directly upstream of a requisite 5'-NGG protospacer adjacent motif (PAM), and Cas9 mediates a double-stranded break (DSB) about 3-base pair upstream of the PAM. In embodiments, the CRISPR nuclease from *Streptococcus aureus* is targeted to genomic DNA by a synthetic guide RNA consisting of a 21-23-nt guide sequence and a scaffold. The guide sequence base-pairs with the DNA target, directly upstream of a requisite 5'-NNGRRT protospacer adjacent motif (PAM), and Cas9 mediates a double-stranded break (DSB) about 3-base pair upstream of the PAM.

[0062] The term “Cas9 variant” refers to proteins that have at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity across the whole sequence or a functional portion of the sequence (e.g. a 50, 100, 150 or 200 continuous amino acid portion) compared to wild-type Cas9 protein and have one or more mutations that increase its binding specificity to PAM compared to wild-type Cas9 protein. Exemplary Cas9 variants are listed in the Table 2 below.

Table 2		
Cas9 Variants	PAM domains	References
Strep pyogenes (Sp) Cas9	NGG	Hsu et al. 2014 Cell
Staph aureus (Sa) Cas9	NNGRRT or NNGRR NNGGGT, NNGAAT, NNGAGT (Zetsche)	Ran et al. 2015 Nature
SpCas9 VQR mutant (D1135V, R1335Q, T1337R)	NGAG>NGAT=NGAA>NGAC NGCG	Kleinstiver et al. 2015 Nature
SpCas9 VRER mutant (D1135V/G1218R/R1335E/T1337R)	NGCG	Kleinstiver et al. 2015 Nature
SpCas9 D1135E	NGG, greater fidelity, less cutting at NAG and NGA sites	Kleinstiver et al. 2015 Nature
eSpCas9 1.1 mutant (K848A/K1003A/R1060A)	NGG	Slaymaker et al. Science 2015
SpCas9 HF1 (Q695A, Q926A, N497A, R661A)	NGG	Kleinstiver et al. 2016 Nature
AsCpfI	TTTN (5' of sgRNA)	Zetsche et al. 2015 Cell

[0063] A “Cfp1” or “Cfp1 protein” as referred to herein includes any of the recombinant or naturally-occurring forms of the Cfp1 (CxxC finger protein 1) endonuclease or variants or homologs thereof that maintain Cfp1 endonuclease enzyme activity (e.g. within at least 50%, 80%, 90%, 95%, 96%, 97%, 98%, 99% or 100% activity compared to Cfp1). In some aspects, the variants or
5 homologs have at least 90%, 95%, 96%, 97%, 98%, 99% or 100% amino acid sequence identity across the whole sequence or a portion of the sequence (e.g. a 50, 100, 150 or 200 continuous amino acid portion) compared to a naturally occurring Cfp1 protein. In embodiments, the Cfp1 protein is substantially identical to the protein identified by the UniProt reference number Q9P0U4 or a variant or homolog having substantial identity thereto.

10 [0064] The term “Class II CRISPR endonuclease” refers to endonucleases that have similar endonuclease activity as Cas9 and participate in a Class II CRISPR system. An example Class II CRISPR system is the type II CRISPR locus from *Streptococcus pyogenes* SF370, which contains a cluster of four genes Cas9, Cas1, Cas2, and Csn1, as well as two non-coding RNA elements, tracrRNA and a characteristic array of repetitive sequences (direct repeats) interspaced by short
15 stretches of non-repetitive sequences (spacers, about 30 bp each). In this system, targeted DNA double-strand break (DSB) may generated in four sequential steps. First, two non-coding RNAs, the pre-crRNA array and tracrRNA, may be transcribed from the CRISPR locus. Second, tracrRNA may hybridize to the direct repeats of pre-crRNA, which is then processed into mature crRNAs containing individual spacer sequences. Third, the mature crRNA:tracrRNA complex may direct
20 Cas9 to the DNA target consisting of the protospacer and the corresponding PAM via heteroduplex formation between the spacer region of the crRNA and the protospacer DNA. Finally, Cas9 may mediate cleavage of target DNA upstream of PAM to create a DSB within the protospacer.

[0065] The term “RNA-guided DNA endonuclease” and the like refer, in the usual and customary sense, to an enzyme that cleave a phosphodiester bond within a DNA polynucleotide chain, wherein
25 the recognition of the phosphodiester bond is facilitated by a separate RNA sequence (for example, a single guide RNA).

[0066] The terms “single guide RNA,” “single guide RNA sequence,” “chimeric RNA,” “chimeric guide RNA,” “guide RNA”, and “synthetic guide RNA” are used interchangeably and refer to the polynucleotide sequence including the crRNA sequence and optionally the tracrRNA
30 sequence. The crRNA sequence includes a guide sequence (i.e., “guide” or “spacer”) and a tracr

mate sequence (i.e., direct repeat(s)). The term “guide sequence” refers to the sequence that specifies the target site.

[0067] In general, a guide sequence is any polynucleotide sequence having sufficient complementarity with a target polynucleotide sequence to hybridize with the target sequence (i.e., a mitochondrial DNA target sequence) and direct sequence-specific binding of a CRISPR complex to the target sequence (i.e., the mitochondrial DNA target sequence). In some embodiments, the degree of complementarity between a guide sequence and its corresponding target sequence, when optimally aligned using a suitable alignment algorithm, is about or more than about 50%, 60%, 75%, 80%, 85%, 90%, 95%, 97.5%, 99%, or more. Optimal alignment may be determined with the use of any suitable algorithm for aligning sequences, non-limiting example of which include the Smith-Waterman algorithm, the Needleman-Wunsch algorithm, algorithms based on the Burrows-Wheeler Transform (e.g. the Burrows Wheeler Aligner), ClustalW, Clustal X, BLAT, Novoalign (Novocraft Technologies, ELAND (Illumina, San Diego, Calif.), SOAP (available at soap.genomics.org.cn), and Maq (available at maq.sourceforge.net). In embodiments, a guide sequence is about or more than about 5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 75, or more nucleotides in length. In embodiments, a guide sequence is less than about 75, 50, 45, 40, 35, 30, 25, 20, 15, 12, or fewer nucleotides in length. The ability of a guide sequence to direct sequence-specific binding of a CRISPR complex to a target sequence may be assessed by any suitable assay. For example, the components of a CRISPR system sufficient to form a CRISPR complex, including the guide sequence to be tested, may be provided to a host cell having the corresponding target sequence, such as by transfection with vectors encoding the components of the CRISPR sequence, followed by an assessment of preferential cleavage within the target sequence, such as by Surveyor assay as described herein. Similarly, cleavage of a target polynucleotide sequence may be evaluated in a test tube by providing the target sequence, components of a CRISPR complex, including the guide sequence to be tested and a control guide sequence different from the test guide sequence, and comparing binding or rate of cleavage at the target sequence between the test and control guide sequence reactions. Other assays are possible, and will occur to those skilled in the art.

[0068] A guide sequence may be selected to target any mitochondrial DNA (mtDNA) target sequence. The term “mitochondrial DNA (mtDNA) target sequence” refers, in the usual and

- customary sense, to a nucleic acid sequence within the mitochondrial genome to which a guide sequence is designed to have complementarity, where hybridization between a target sequence and a guide sequence promotes the formation of a CRISPR complex. Full complementarity is not necessarily required, provided there is sufficient complementarity to cause hybridization and promote formation of a CRISPR complex. A guide sequence (spacer) may comprise any polynucleotide, such as DNA or RNA polynucleotides. In embodiments, a guide sequence (spacer) includes a nucleic acid sequence of one of SEQ ID Nos: 55-105. A target sequence may comprise any polynucleotide, such as DNA or RNA polynucleotides. In embodiments, an mtDNA target sequence includes one or more point mutations, one or more deletions, or any combination thereof.
- 10 In embodiments, an mtDNA target sequence includes a nucleic acid sequence of one of SEQ ID Nos:8-22 or a fragment thereof. In embodiments, the mtDNA target sequence includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to one of SEQ ID Nos:8-22 or a fragment thereof.

Table 3		
mtDNA targets	mtDNA mutation sequence	target/spacer RNA
14787 cytochrome b (cyt b) deletion 4bp (TTAA)	CCAATGACCCCAATACGCAAAA TTAACCCCCTAATAAAACTAA _CCACTCATTCATCGACCTCCC CACCCCATCCAACATCTCCGC (SEQ ID NO: 8)	1. GAGTGGTTAGTTTTATTAGG (SEQ ID NO: 55) 2. AATGAGTGGTTAGTTTTATTA (SEQ ID NO: 56)
3243A>G	CAGGGTTTGTTAAGATGGCAGg GCCCGGTAATCGCATAAAACTT AA (SEQ ID NO: 9)	1. CAGGGTTTGTTAAGATGGCA (SEQ ID NO: 57) 2. ACAGGGTTTGTTAAGATGGC (SEQ ID NO: 58)
3271T>C	CCCGGTAATCGCATAAAACTTA AAACcTTACAGTCAGAGGTTCA ATTCCTCTTCTT (SEQ ID NO: 10)	1. GAATTGAACCTCTGACTGTA (SEQ ID NO: 59) 2. ACTTAAACCTTACAGTCAG (SEQ ID NO: 60)
8344A>G	GCATTAACCTTTTAAGTTAAAG ATTAAGAGAgCCAACACCTCTC TACAGTGAAATGCCCAACTA (SEQ ID NO: 11)	1. CTTTAAAGTTAAAGATTAAG (SEQ ID NO: 61) 2. TTTAAGTTAAAGATTAAGAG (SEQ ID NO: 62)
8356T>C	CCTTTAAGTTAAAGATTAAGA GAACCAACACCTCTcTACAGTG AAATGCCCAACTAAATAC (SEQ ID NO: 12)	1. TGGGGCATTTCAGTGTAGAG (SEQ ID NO: 63) 2. TAGTTGGGGCATTTCAGTGT (SEQ ID NO: 64)
8993T>G	CCATCAGCCTACTCATTCAACC AATAGCCCgGGCCGTACGCCTA ACCGCTAACATTACTGCAG (SEQ ID NO: 13)	1. AGGCGTACGGCCCGGGCTAT (SEQ ID NO: 65) 2. TACTCATTCAACCAATAGCC (SEQ ID NO: 66) 3. AGCGGTTAGGCGTACGGCCC (SEQ ID NO: 67)

		NO: 67) 4. GGC GTACGGCCCCGGGCTATTGG (SEQ ID NO: 68)
8993T>C	CCATCAGCCTACTCATTCAACC AATAGCCCcGGCCGTACGCCTA ACCGCTAACATTACTGCAG (SEQ ID NO: 14)	1. AGGCGTACGGCCGGGGCTAT (SEQ ID NO: 69) 2. TTAGCGGTTAGGCGTACGGC (SEQ ID NO: 70) 3. AGCGGTTAGGCGTACGGCCG (SEQ ID NO: 71) 4. GCGTACGGCCGGGGCTATTGG (SEQ ID NO: 72)
3460G>A	GGGCTACTACAACCCTTCGCTG ACaCCATAAACTCTTCACCAA AGAGCCCCT (SEQ ID NO: 15)	1. GAGTTTTATGGTGT CAGCGA (SEQ ID NO: 73) 2. GAAGAGTTTTATGGTGT CAGCG (SEQ ID NO: 74) 3. GTGAAGAGTTTTATGGTGT CAGC (SEQ ID NO: 75) 4. TGGTGT CAGCGAAGGGTTGTAGT (SEQ ID NO: 76) 5. GTGAAGAGTTTTATGGTGT C (SEQ ID NO: 77)
11778G>A	GCCTAGCAA ACTCAA ACTACGA ACGCACTCACAGTcCATCATA ATCCTCTCTCAAGGACTTCAAA CTCTACTCCC (SEQ ID NO: 16)	1. GAGAGAGGATTATGATGTGACT (SEQ ID NO: 78) 2. CACATCATAATCCTCTCTCA (SEQ ID NO: 79) 3. GATTATGATGTGACTGTGAG (SEQ ID NO: 80) 4. GAGGATTATGATGTGACTG (SEQ ID NO: 81)
14484T>C	GCCATCGCTGTAGTATATCCAA AGACAACCAcCATTCCCCCTAA ATAAATTAAAAAACTATTAA CC (SEQ ID NO: 17)	1. TTAAATTTATTTAGGGGGAA (SEQ ID NO: 82) 2. AATTTATTTAGGGGGAAATGG (SEQ ID NO: 83) 3. GGGGGAATGGTGGTTGTCTT (SEQ ID NO: 84) 4. TAGGGGGAATGGTGGTTGTCT (SEQ ID NO: 85) 5. ATTTATTTAGGGGGAAATGGTGGT (SEQ ID NO: 86) 6. GGGGGAATGGTGGTTGTCTTTGG (SEQ ID NO: 87) 7. TTAAATTTATTTAGGGGGAAATGG (SEQ ID NO: 88)
14709T>C	ATTCTCGCACGGACTACAACCA CGACCAATGATAcGAAAAACCA TCGTTGTATTTCAACTACAAGA (SEQ ID NO: 18)	1. ACGATGGTTTTTCGTATCAT (SEQ ID NO: 89) 2. GTATCATTGGTCGTGGTTGTAGT (SEQ ID NO: 90) 3. TCGTATCATTGGTCGTGGTTGTA (SEQ ID NO: 91) 4. GTTTTTCGTATCATTGGTCG (SEQ ID

		NO: 92)
1555A>G	CCCCTACGCATTTATATAGAGG AGACAAGTCGTAACATGGTAAG (SEQ ID NO: 19)	1. CCCCTACGCATTTATATAGA (SEQ ID NO: 93) 2. AGAGGAGACAAGTCGTAACA (SEQ ID NO: 94) 3. TGTCTCCTCTATATAAATGCGT (SEQ ID NO: 95) 4. TATAGAGGAGACAAGTCGTAACA (SEQ ID NO: 96)
10158T>C	GACTACCACAACCTCAACGGCTA CATAGAAAAAaCCACCCCTTAC GAGTGCGGCTTCGACCC (SEQ ID NO: 20)	1. AAACCCACCCCTTACGAGTG (SEQ ID NO: 97) 2. AAGCCGCACTCGTAAGGGGT (SEQ ID NO: 98) 3. CATAGAAAAACCCACCCCTT (SEQ ID NO: 99)
10191T>C	CCCCTTACGAGTGCGGCTTCGA CCCTATAcCCCCGCGCGTCC CTTCTCCATAAAATTCTTCTTA G (SEQ ID NO: 21)	1. GGACGCGGGCGGGGGGTATA (SEQ ID NO: 100) 2. GAGAAAGGGACGCGGGCGGG (SEQ ID NO: 101) 3. GAAAGGGACGCGGGCGGGGGTA (SEQ ID NO: 102)
6930G>A	GAAATGATCTGCTGCAGTGCTC TGAGCCCTAaGATTCATCTTTCT TTTCACCGTAGGTGGCCTGACT GGC (SEQ ID NO: 22)	1. GAAAAGAAAGATGAATCTT (SEQ ID NO: 103) 2. AAGATGAATCTTAGGGCTC (SEQ ID NO: 104) 3. GATGAATCTTAGGGCTCAG (SEQ ID NO: 105)

[0069] In general, a tracr mate sequence includes any sequence that has sufficient complementarity with a tracr sequence (i.e., a tracrRNA sequence) to promote one or more of: (1) excision of a guide sequence flanked by tracr mate sequences in a cell containing the corresponding tracr sequence; and

5 (2) formation of a CRISPR complex at a target sequence, wherein the CRISPR complex comprises the tracr mate sequence hybridized to the tracr sequence. In general, degree of complementarity is with reference to the optimal alignment of the tracr mate sequence and tracr sequence, along the length of the shorter of the two sequences. Optimal alignment may be determined by any suitable alignment algorithm, and may further account for secondary structures, such as self-complementarity

10 within either the tracr sequence or tracr mate sequence. In some embodiments, the degree of complementarity between the tracr sequence and tracr mate sequence along the length of the shorter of the two when optimally aligned is about or more than about 25%, 30%, 40%, 50%, 60%, 70%, 80%, 90%, 95%, 97.5%, 99%, or higher. In some embodiments, the tracr sequence is about or more than about 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 40, 50, or more nucleotides

in length. In some embodiments, the tracr sequence and tracr mate sequence are contained within a single transcript, such that hybridization between the two produces a transcript having a secondary structure, such as a hairpin.

[0070] Without wishing to be bound by theory, the tracr sequence, which may comprise or consist of all or a portion of a wild-type tracr sequence (e.g. about or more than about 20, 26, 32, 45, 48, 54, 63, 67, 85, or more nucleotides of a wild-type tracr sequence), may also form part of a CRISPR complex, such as by hybridization along at least a portion of the tracr sequence to all or a portion of a tracr mate sequence that is operably linked to the guide sequence. In some embodiments, the tracr sequence has sufficient complementarity to a tracr mate sequence to hybridize and participate in formation of a CRISPR complex. As with the target sequence, it is believed that complete complementarity is not needed, provided there is sufficient to be functional. In some embodiments, the tracr sequence has at least 50%, 60%, 70%, 80%, 90%, 95% or 99% of sequence complementarity along the length of the tracr mate sequence when optimally aligned. Where the tracrRNA sequence is less than 100 (99 or less) nucleotides in length the sequence is one of 99, 98, 97, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, or 20 nucleotides in length.

[0071] In embodiments, the tracrRNA sequence for *S. pyogenes* is:
 5'-gttggaaccattcaaaacagcatagcaagttaaataaggctagtcggttatcaacttgaaaaagtggcaccgagtcggtgctttt-3' (SEQ ID NO: 23). In embodiments, the tracrRNA sequence may have 80%, 85%, 90%, 91 %, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% sequence identity to SEQ ID NO: 23.

[0072] In embodiments, the crRNA with or without tracrRNA sequence for exemplary Class II CRISPR endonucleases may include:

25

Table 4. crRNA ± tracrRNA of RNA-mediated endonuclease	
SpCas9	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTC CGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGC (SEQ ID NO: 106)
SaCas9	GTTTTAGTACTCTGGAAACAGAATCTACTAAAACAAGGCA AAATGCCGTGTTTATCTCGTCAACTTGTGGCGAGA (SEQ ID NO: 107)
Cpf1	TAATTTCTACTCTTGTAGAT (SEQ ID NO: 108)

[0073] The term “transfecting,” “transfection” and the like refer, in the usual and customary sense, to the process of introducing nucleic acids into cells.

- 5 **[0074]** The compositions described herein can be purified. Purified compositions are at least about 60% by weight (dry weight) the compound of interest. Preferably, the preparation is at least about 75%, more preferably at least about 90%, and most preferably at least about 99% or higher by weight the compound of interest. Purity is measured by any appropriate standard method, for example, by High-performance liquid chromatography, polyacrylamide gel electrophoresis.
- 10 **[0075]** A “cell” as used herein, refers to a cell carrying out metabolic or other function sufficient to preserve or replicate its genomic DNA. A cell can be identified by well-known methods in the art including, for example, presence of an intact membrane, staining by a particular dye, ability to produce progeny or, in the case of a gamete, ability to combine with a second gamete to produce a viable offspring. Cells may include prokaryotic and eukaryotic cells. Prokaryotic cells include but
- 15 are not limited to bacteria. Eukaryotic cells include but are not limited to yeast cells and cells derived from plants and animals, for example mammalian, insect (*e.g.*, spodoptera) and human cells.

- [0076]** As used herein, the term “vector” refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a “plasmid”, which refers to a linear or circular double stranded DNA loop into which additional DNA segments can be ligated.
- 20 Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (*e.g.*, bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (*e.g.*, non episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the

host genome. Moreover, certain vectors are capable of directing the expression of genes to which they are operatively linked. Such vectors are referred to herein as “expression vectors”. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids. In the present specification, “plasmid” and “vector” can be used interchangeably as the plasmid is the most commonly used form of vector. However, the invention is intended to include such other forms of expression vectors, such as viral vectors (*e.g.*, replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions. Additionally, some viral vectors are capable of targeting a particular cells type either specifically or non-specifically. Replication-incompetent viral vectors or replication-defective viral vectors refer to viral vectors that are capable of infecting their target cells and delivering their viral payload, but then fail to continue the typical lytic pathway that leads to cell lysis and death.

[0077] A “pharmaceutical composition” is a formulation containing the nucleic acids described herein in a form suitable for administration to a subject. In embodiments, the pharmaceutical composition is in bulk or in unit dosage form. The unit dosage form is any of a variety of forms, including, for example, a capsule, an IV bag, a tablet, a single pump on an aerosol inhaler or a vial. The quantity of active ingredient (*e.g.*, a formulation of the disclosed nucleic acid) in a unit dose of composition is an effective amount and is varied according to the particular treatment involved. One skilled in the art will appreciate that it is sometimes necessary to make routine variations to the dosage depending on the age and condition of the patient. The dosage will also depend on the route of administration. A variety of routes are contemplated, including oral, pulmonary, rectal, parenteral, transdermal, subcutaneous, intravenous, intramuscular, intraperitoneal, inhalational, buccal, sublingual, intrapleural, intrathecal, intranasal, and the like. Dosage forms for the topical or transdermal administration of a compound of this invention include powders, sprays, ointments, pastes, creams, lotions, gels, solutions, patches and inhalants. In embodiments, the active nucleic acid is mixed under sterile conditions with a pharmaceutically acceptable carrier, and with any preservatives, buffers, or propellants that are required.

[0078] As used herein, the phrase “pharmaceutically acceptable” refers to those compounds, anions, cations, materials, compositions, carriers, and/or dosage forms which are, within the scope of sound medical judgment, suitable for use in contact with the tissues of human beings and animals

without excessive toxicity, irritation, allergic response, or other problem or complication, commensurate with a reasonable benefit/risk ratio.

[0079] “Pharmaceutically acceptable excipient” means an excipient that is useful in preparing a pharmaceutical composition that is generally safe, non-toxic and neither biologically nor otherwise

undesirable, and includes excipient that is acceptable for veterinary use as well as human pharmaceutical use. A “pharmaceutically acceptable excipient” as used in the specification and claims includes both one and more than one such excipient. A thorough discussion of pharmaceutically acceptable excipients is available in REMINGTON'S PHARMACEUTICAL SCIENCES (Mack Pub. Co., N.J. 1991). Pharmaceutically acceptable excipients in therapeutic compositions may contain liquids such as water, saline, glycerol and ethanol. Additionally, auxiliary substances, such as wetting or emulsifying agents, pH buffering substances, and the like, may be present in such vehicles.

[0080] A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, *e.g.*, intravenous, intradermal, subcutaneous, oral (*e.g.*, inhalation), transdermal (topical), and transmucosal administration.

[0081] Formulations suitable for oral administration can consist of (a) liquid solutions, such as an effective amount of the packaged nucleic acid suspended in diluents, such as water, saline or PEG 400; (b) capsules, sachets or tablets, each containing a predetermined amount of the active ingredient, as liquids, solids, granules or gelatin; (c) suspensions in an appropriate liquid; and (d) suitable emulsions. Tablet forms can include one or more of lactose, sucrose, mannitol, sorbitol, calcium phosphates, corn starch, potato starch, microcrystalline cellulose, gelatin, colloidal silicon dioxide, talc, magnesium stearate, stearic acid, and other excipients, colorants, fillers, binders, diluents, buffering agents, moistening agents, preservatives, flavoring agents, dyes, disintegrating agents, and pharmaceutically compatible carriers. Lozenge forms can comprise the active ingredient in a flavor, *e.g.*, sucrose, as well as pastilles comprising the active ingredient in an inert base, such as gelatin and glycerin or sucrose and acacia emulsions, gels, and the like containing, in addition to the active ingredient, carriers known in the art.

[0082] Pharmaceutical compositions can also include large, slowly metabolized macromolecules such as proteins, polysaccharides such as chitosan, polylactic acids, polyglycolic acids and copolymers (such as latex functionalized sepharose(TM), agarose, cellulose, and the like), polymeric amino acids, amino acid copolymers, and lipid aggregates (such as oil droplets or liposomes).

5 Additionally, these carriers can function as immunostimulating agents (*i.e.*, adjuvants).

[0083] Suitable formulations for rectal administration include, for example, suppositories, which consist of the packaged nucleic acid with a suppository base. Suitable suppository bases include natural or synthetic triglycerides or paraffin hydrocarbons. In addition, it is also possible to use gelatin rectal capsules which consist of a combination of the compound of choice with a base,
10 including, for example, liquid triglycerides, polyethylene glycols, and paraffin hydrocarbons.

[0084] Formulations suitable for parenteral administration, such as, for example, by intraarticular (in the joints), intravenous, intramuscular, intratumoral, intradermal, intraperitoneal, and subcutaneous routes, include aqueous and non-aqueous, isotonic sterile injection solutions, which can contain antioxidants, buffers, bacteriostats, and solutes that render the formulation isotonic with
15 the blood of the intended recipient, and aqueous and non-aqueous sterile suspensions that can include suspending agents, solubilizers, thickening agents, stabilizers, and preservatives. In the practice of this invention, compositions can be administered, for example, by intravenous infusion, orally, topically, intraperitoneally, intravesically or intrathecally. Parenteral administration, oral administration, and intravenous administration are the preferred methods of administration. The
20 formulations of compounds can be presented in unit-dose or multi-dose sealed containers, such as ampules and vials.

[0085] Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial
25 agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates, and agents for the adjustment of tonicity such as sodium chloride or dextrose. The pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made
30 of glass or plastic.

[0086] A pharmaceutical composition of the invention can be administered to a subject in many of the well-known methods currently used for chemotherapeutic treatment. For example, for treatment of cancers, a composition of the invention may be injected directly into tumors, injected into the blood stream or body cavities or taken orally or applied through the skin with patches. The dose
5 chosen should be sufficient to constitute effective treatment but not so high as to cause unacceptable side effects. The state of the disease condition (*e.g.*, cancer, precancer, and the like) and the health of the patient should preferably be closely monitored during and for a reasonable period after treatment.

[0087] The terms “subject,” “individual,” and “patient” are used interchangeably herein to refer to
10 a vertebrate, preferably a mammal, more preferably a human. The mammal can be *e.g.*, a human or appropriate non-human mammal, such as primate, mouse, rat, dog, cat, cow, horse, goat, camel, sheep or a pig. The subject can also be a bird or fowl. Tissues, cells and their progeny of a biological entity obtained *in vivo* or cultured *in vitro* are also encompassed. As used herein, a “subject in need thereof” or “a patient” may be a subject having a mitochondria disease.

[0088] As used herein, “treatment” or “treating,” or “palliating” or “ameliorating” are used
15 interchangeably herein. These terms refer to an approach for obtaining beneficial or desired results including but not limited to therapeutic benefit and/or a prophylactic benefit. By therapeutic benefit is meant eradication or amelioration of the underlying disorder being treated. Also, a therapeutic benefit is achieved with the eradication or amelioration of one or more of the physiological
20 symptoms associated with the underlying disorder such that an improvement is observed in the patient, notwithstanding that the patient may still be afflicted with the underlying disorder. For prophylactic benefit, the compositions may be administered to a patient at risk of developing a particular disease, or to a patient reporting one or more of the physiological symptoms of a disease, even though a diagnosis of this disease may not have been made. Treatment includes preventing the
25 disease, that is, causing the clinical symptoms of the disease not to develop by administration of a protective composition prior to the induction of the disease; suppressing the disease, that is, causing the clinical symptoms of the disease not to develop by administration of a protective composition after the inductive event but prior to the clinical appearance or reappearance of the disease;
inhibiting the disease, that is, arresting the development of clinical symptoms by administration of a
30 protective composition after their initial appearance; preventing re-occurring of the disease and/or

relieving the disease, that is, causing the regression of clinical symptoms by administration of a protective composition after their initial appearance.

[0089] An “effective amount” is an amount sufficient to accomplish a stated purpose (e.g. achieve the effect for which it is administered, treat a disease, reduce enzyme activity, reduce one or more symptoms of a disease or condition, reduce viral replication in a cell). An example of an “effective amount” is an amount sufficient to contribute to the treatment, prevention, or reduction of a symptom or symptoms of a disease, which could also be referred to as a “therapeutically effective amount.” A “reduction” of a symptom or symptoms (and grammatical equivalents of this phrase) means decreasing of the severity or frequency of the symptom(s), or elimination of the symptom(s).

A “prophylactically effective amount” of a drug is an amount of a drug that, when administered to a subject, will have the intended prophylactic effect, e.g., preventing or delaying the onset (or reoccurrence) of an injury, disease, pathology or condition, or reducing the likelihood of the onset (or reoccurrence) of an injury, disease, pathology, or condition, or their symptoms. The full prophylactic effect does not necessarily occur by administration of one dose, and may occur only after administration of a series of doses. Thus, a prophylactically effective amount may be administered in one or more administrations. An “activity decreasing amount,” as used herein, refers to an amount of antagonist required to decrease the activity of an enzyme or protein relative to the absence of the antagonist. A “function disrupting amount,” as used herein, refers to the amount of antagonist required to disrupt the function of an enzyme or protein relative to the absence of the antagonist. Guidance can be found in the literature for appropriate dosages for given classes of pharmaceutical products. For example, for the given parameter, an effective amount will show an increase or decrease of at least 5%, 10%, 15%, 20%, 25%, 40%, 50%, 60%, 75%, 80%, 90%, or at least 100%. Efficacy can also be expressed as “-fold” increase or decrease. For example, a therapeutically effective amount can have at least a 1.2-fold, 1.5-fold, 2-fold, 5-fold, or more effect over a control. The exact amounts will depend on the purpose of the treatment, and will be ascertainable by one skilled in the art using known techniques (*see, e.g., Lieberman, Pharmaceutical Dosage Forms* (vols. 1-3, 1992); Lloyd, *The Art, Science and Technology of Pharmaceutical Compounding* (1999); Pickar, *Dosage Calculations* (1999); and *Remington: The Science and Practice of Pharmacy*, 20th Edition, 2003, Gennaro, Ed., Lippincott, Williams & Wilkins).

[0090] As used herein, the term “administering” means oral administration, administration as a suppository, topical contact, intravenous, intraperitoneal, intramuscular, intralesional, intrathecal, intranasal or subcutaneous administration, or the implantation of a slow-release device, *e.g.*, a mini-osmotic pump, to a subject. Administration is by any route, including parenteral and transmucosal
 5 (*e.g.*, buccal, sublingual, palatal, gingival, nasal, vaginal, rectal, or transdermal).

[0091] As used herein, “mitochondrial disorders” related to disorders which are due to abnormal mitochondria such as for example, a mitochondrial genetic mutation, enzyme pathways etc.

Examples of disorders include and are not limited to: loss of motor control, muscle weakness and pain, gastro-intestinal disorders and swallowing difficulties, poor growth, cardiac disease, liver
 10 disease, diabetes, respiratory complications, seizures, visual/hearing problems, lactic acidosis, developmental delays and susceptibility to infection. The mitochondrial abnormalities give rise to

“mitochondrial diseases” which include, but not limited to: AD: Alzheimer's Disease; ADPD:

Alzheimer's Disease and Parkinson's Disease; AMDF: Ataxia, Myoclonus and Deafness CIPO:

Chronic Intestinal Pseudoobstruction with myopathy and Ophthalmoplegia; CPEO: Chronic
 15 Progressive External Ophthalmoplegia; DEAF: Maternally inherited DEAFness or aminoglycoside-

induced DEAFness; DEMCHO: Dementia and Chorea; DMDF: Diabetes Mellitus & Deafness;
 Exercise Intolerance; ESOC: Epilepsy, Strokes, Optic atrophy, & Cognitive decline; FBSN: Familial

Bilateral Striatal Necrosis; FICP: Fatal Infantile Cardiomyopathy Plus, a MELAS-associated
 cardiomyopathy; GER: Gastrointestinal Reflux; KSS Kearns Sayre Syndrome LDYT: Leber's
 20 hereditary optic neuropathy and DYsTonia; LHON: Leber Hereditary Optic Neuropathy; LIMM:

Lethal Infantile Mitochondrial Myopathy; MDM: Myopathy and Diabetes Mellitus; MELAS:
 Mitochondrial Encephalomyopathy, Lactic Acidosis, and Stroke-like episodes; MEPR: Myoclonic
 Epilepsy and Psychomotor Regression; MERME: MERRF/MELAS overlap disease; MERRF:

Myoclonic Epilepsy and Ragged Red Muscle Fibers; MHCM: Maternally Inherited Hypertrophic
 25 CardioMyopathy; MICM: Maternally Inherited Cardiomyopathy; MILS: Maternally Inherited Leigh
 Syndrome; Mitochondrial Encephalocardiomyopathy; Mitochondrial Encephalomyopathy; MM:

Mitochondrial Myopathy; MMC: Maternal Myopathy and Cardiomyopathy; Multisystem
 Mitochondrial Disorder (myopathy, encephalopathy, blindness, hearing loss, peripheral neuropathy);
 NARP: Neurogenic muscle weakness, Ataxia, and Retinitis Pigmentosa; alternate phenotype at this
 30 locus is reported as Leigh Disease; NIDDM: Non-Insulin Dependent Diabetes Mellitus; PEM:

Progressive Encephalopathy; PME: Progressive Myoclonus Epilepsy; RTT: Rett Syndrome; SIDS: Sudden Infant Death Syndrome.

[0092] In embodiments, a mitochondrial disorder is selected from the group consisting of Myoclonic Epilepsy with Ragged Red Fibers (MERRF); Mitochondrial Myopathy, Encephalopathy, Lactacidosis, and Stroke (MELAS); Maternally Inherited Diabetes and Deafness (MIDD); Leber's Hereditary Optic Neuropathy (LHON); chronic progressive external ophthalmoplegia (CPEO); Leigh Disease; Kearns-Sayre Syndrome (KSS); Friedreich's Ataxia (FRDA); Co-Enzyme Q10 (CoQ10) Deficiency; Complex I Deficiency; Complex II Deficiency; Complex III Deficiency; Complex IV Deficiency; Complex V Deficiency; other myopathies; cardiomyopathy; encephalomyopathy; renal tubular acidosis; neurodegenerative diseases; Parkinson's disease; Alzheimer's disease; amyotrophic lateral sclerosis (ALS); motor neuron diseases; hearing and balance impairments; or other neurological disorders; epilepsy; genetic diseases; Huntington's Disease; mood disorders; nucleoside reverse transcriptase inhibitors (NRTI) treatment; HIV-associated neuropathy; schizophrenia; bipolar disorder; age-associated diseases; cerebral vascular diseases; macular degeneration; diabetes; and cancer.

[0093] "Contacting" is used in accordance with its plain ordinary meaning and refers to the process of allowing at least two distinct species (e.g. chemical compounds including biomolecules or cells) to become sufficiently proximal to react, interact or physically touch. It should be appreciated, however, that the resulting reaction product can be produced directly from a reaction between the added reagents or from an intermediate from one or more of the added reagents which can be produced in the reaction mixture.

[0094] A "control" sample or value refers to a sample that serves as a reference, usually a known reference, for comparison to a test sample. For example, a test sample can be taken from a test condition, e.g., in the presence of a test compound, and compared to samples from known conditions, e.g., in the absence of the test compound (negative control), or in the presence of a known compound (positive control). A control can also represent an average value gathered from a number of tests or results. One of skill in the art will recognize that controls can be designed for assessment of any number of parameters. For example, a control can be devised to compare therapeutic benefit based on pharmacological data (e.g., half-life) or therapeutic measures (e.g., comparison of side effects). One of skill in the art will understand which standard controls are most

appropriate in a given situation and be able to analyze data based on comparisons to standard control values. Standard controls are also valuable for determining the significance (e.g. statistical significance) of data. For example, if values for a given parameter are widely variant in standard controls, variation in test samples will not be considered as significant.

5 **[0095]** As used herein, the term “about” means a range of values including the specified value, which a person of ordinary skill in the art would consider reasonably similar to the specified value. In embodiments, the term “about” means within a standard deviation using measurements generally acceptable in the art. In embodiments, about means a range extending to +/- 10% of the specified value. In embodiments, about means the specified value.

10 **II. Compositions**

[0096] In a first aspect, there is provided a composition including a delivery vehicle and a protein including a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme, wherein the protein is bound to the delivery vehicle. The term “delivery vehicle” or “carrier” refers to any support structure that brings about the transfer of a
15 component of genetic material or a protein. Genetic material includes but is not limited to DNA, RNA or fragments thereof and proteins or polypeptides comprise amino acids and include but are not limited to antigens, antibodies, ligands, receptors or fragments thereof. Delivery vehicles include but are not limited to vectors such as viruses (examples include but are not limited to retroviruses, adenoviruses, adeno-associated viruses, pseudotyped viruses, replication competent
20 viruses, herpes simplex virus), virus capsids, liposomes or liposomal vesicles, lipoplexes, polyplexes, dendrimers, macrophages, artificial chromosomes, nanoparticles, polymers and also hybrid particles, examples of which include virosomes. Delivery vehicles may have multiple surfaces and compartments for attachment and storage of components. These include but are not limited to outer surfaces and inner compartments.

25 **[0097]** In embodiments, the delivery vehicle is a nanoparticle or a lipid particle or a viral vector. Any nanoparticles known for protein or nucleic acid delivery can be used for the invention described herein. Nanoparticles are particles between 1 and 100 nanometers in size. Recent dramatic advances in nanotechnology have led to the development of a variety of nanoparticles (NPs) that provide valuable tools. Numerous nanomaterials such as polymers, liposomes, protein based NPs

and inorganic NPs have been developed and a variety of particles are currently being evaluated in clinical studies with promising initial results; and some liposomal NPs are approved by the FDA. One of the major advantages of using these NPs is that they offer targeted tissue/site delivery. Their small size allows NPs to escape through blood vessels at the target site through the leaky vascular structure (Enhanced permeability and retention effect). In addition to this passive mechanism, a variety of targeting moieties can be attached to NPs to confer active targeting capability. Exemplary nanoparticles that can be used for delivering compositions described herein include, but are not limited to, solid nanoparticles (e.g., metals such as silver, gold, iron, titanium), non-metal, lipid-based solids (e.g., liposome), polymers (e.g., polyethylenimine, dendrimer), suspensions of nanoparticles, or combinations thereof (e.g., polyethylenimine-liposome, dendrisome). Any compositions described herein (such as Mito-Cas9, mito-Cpf1, or other mito-RNA guided nucleases (mito-RGN)) may be delivered in nanoparticle complexes in the form of protein, DNA, or mRNA. Additional information about nanoparticles that can be used by the compositions described herein can be found in Coelho et al., N Engl J Med 2013;369:819-29, Tabernero et al., Cancer Discovery, April 2013, Vol. 3, No. 4, pages 363-470, Zhang et al., WO2015089419 A2, and Zuris JA et al., Nat Biotechnol. 2015;33(1):73-80, each of which is incorporated herein by reference.

[0098] In embodiments, the vector is a replication-incompetent viral vector. For example, the replication-incompetent viral vector is a replication-incompetent DNA viral vector (including, but is not limited to, adenoviruses, adeno-associated viruses). For example, the replication-incompetent viral vector is a replication-incompetent RNA viral vector (including, but is not limited to, replication defective retroviruses, lentiviruses, and rabies viruses).

[0099] In embodiments, the delivery vehicle is a lipid particle-a particle having lipid as a component, such as liposomes or liposomal vesicles or lipoplexes. Liposomes, also known as vesicles, are generally composed of phospholipids and other lipid components such as cholesterol. They can function as carriers whose essential structural feature is a bipolar lipid membrane which envelops an aqueous core volume in which pharmacological agents are solubilized and therefore encapsulated. Various lipid formulations and methods for their preparation have been described for the delivery of pharmaceutically active agents to a host. For example, Geho and Lau in U.S. Pat. No. 4,603,044 describe a targeted liposomal delivery system for delivery of a drug to the hepatobiliary receptors of the liver. The system is composed of a drug or diagnostic agent

encapsulated in or associated with lipid membrane structures in the form of vesicles or liposomes, and a molecule having a fatty substituent attached to the vesicle wall and a target substituent which is a biliary attracted chemical, such as a substituted iminodiacetate complex. The system is particularly useful for the delivery of insulin and serotonin in the treatment of Types I and II diabetes, respectively. Several cationic lipid reagents have become commercially available for transfecting eukaryotic cells. These examples include Lipofectin® (DOTMA:DOPE) (Invitrogen, Carlsbad, CA), LipofectAmine™ (DOSPA:DOPE)(Invitrogen), LipofectAmine2000™ (Invitrogen), LipofectAmine 3000™ (Invitrogen), Lipofectamine RNAiMax™ (Invitrogen), Lipofectamine LTX™ (Thermo Fisher Scientific), Fugene®, Transfectam® (DOGS), Effectene®, DC-Chol. US Patent Publication No. 20050019923 involves cationic dendrimers for delivering bioactive molecules, such as polynucleotide molecules, peptides and polypeptides and/or pharmaceutical agents, to a mammalian body, given the low toxicity and targeting specificity. Other derivatives of cationic dendrimer mentioned in Bioactive Polymers, US published application 20080267903, may also be suitable delivery vehicles for mitoCas9 gene therapy.

[0100] Various polymeric formulations of biologically active agents and methods for their preparation have been described. U.S. Patent Nos. 3,773,919, 3,991,776, 4,076,779, 4,093,709, 4,118,470, 4,131,648, 4,138,344, 4,293,539 and 4,675,189, inter alia, disclose the preparation and use of biocompatible, biodegradable polymers, such as poly (lactic acid), poly(glycolic acid), copolymers of glycolic and lactic acids, poly (o-hydroxycarboxylic acid), polylactones, polyacetals, polyorthoesters and polyorthocarbonates, for the encapsulation of drugs and medicaments. These polymers mechanically entrap the active constituents and later provide controlled release of the active ingredient via polymer dissolution or degradation. Certain condensation polymers formed from divinyl ethers and polyols are described in Polymer Letters, 18, 293 (1980). Polymers have proven to be successful controlled-release drug delivery devices.

[0101] More information about liposomal constructs or polymeric constructs that can be used for the present invention can be found at Schwendener RA et al., Ther Adv Vaccines. 2014 Nov; 2(6): 159–182; Li Y et al., J Gene 2011, Med 13: 60–72; Pichon C et al., Methods Mol Biol 2013 969: 247–274; McNamara MA et al., J Immunol Res. 2015; 2015: 794528; Sayour E. J. et al., *Journal for Immunotherapy of Cancer*. 2015;3, article 13; Bettinger T. et al, *Current Opinion in Molecular Therapeutics*. 2001;3(2):116–124; Lu D. et al., *Cancer Gene Therapy*. 1994;1(4):245–252;

Wasungu L. et al., *Journal of Controlled Release*. 2006;116(2):255–264; Little S. et al., *Proceedings of the National Academy of Sciences of the United States of America*. 2004;101(26):9534–9539; Phua K. et al., *Journal of Controlled Release*. 2013;166(3):227–233; Su X et al., *Molecular Pharmaceutics*. 2011;8(3):774–787; Phua K. K. L. et al., *Nanoscale*. 2014;6(14):7715–7729; Phua K. K. L. et al., *Scientific Reports*. 2014;4, article 5128.

[0102] In embodiments, the protein is encapsulated within said delivery vehicle. Encapsulation can be carried out by any methods known in the art.

[0103] In embodiments, the mitochondrial localization amino acid sequence is N-terminal to the RNA-guided DNA endonuclease enzyme.

[0104] In embodiments, the mitochondrial localization amino acid sequence is the mitochondrial localization amino acid sequence of cytochrome c oxidase subunit VIII (Cox8) sequence.

[0105] In embodiments, the RNA-guided DNA endonuclease enzyme is Cas9 or Cpf1 or a Class II CRISPR endonuclease.

[0106] In embodiments, the RNA-guided DNA endonuclease enzyme is Cas9 (aka *Strep pyogenes* Cas9), containing the amino acid sequence of SEQ ID NO: 24 or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 24 (or appropriate portion thereof).

DKKYSIGLDIGTNSVGAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL
 QEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRILIYLAHMIK
 FRGHFLIEGDLNPDNSDVKLFQVLQTYNQLFEENPINASGVDAKILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKQLSKDQYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIFKPILEKMDGTEELLVKLNREDLLRKQRTFD
 NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGAS
 AQSFIERMTNFDKNLPNEKVLPHSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY
 FKKIECFDSVEISGVEDRFNASLGTYHDLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKYAHLFDDKVMK
 QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDFANRNFQMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
 PAIKKGILQTVKVVDLVKVMGRHKPENIVIEARENQTTQKGQKNSRERMKRIEEDIKELGSQILKEHPVENTQLQNEKL
 YLYYLQNGRDMYVDQELDINRLSDYDQVDHIVPQSFLKDDSIDNKVLTNRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLI
 TQRFKDNLTKAERGGLSELDKAGFIKRLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KVRINNYHHAHDAYLNAVVGTAIIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLAN
 GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD
 SPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEFKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLAS
 AGELQKGNELALPSKYVNFYLAHYEKLKGSPEQKQKLFVEQHKHYLDEIIIEQISEFSKRVIADANLDKVL SAYNKH
 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDTLHQSI TGLYETRIDLSQLGGD (SEQ ID
 NO: 24)

[0107] In embodiments, the RNA-guided DNA endonuclease enzyme is Cpf1 (aka AsCpf1), containing the amino acid sequence of SEQ ID NO: 25 or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 25 (or appropriate portion thereof).

5 TQFEGFTNLYQVSKTLRFELIPQGKTLKHIQEQGFIEEDKARNDDHYKELKPIIDRIYKTYADQCLQLVQLDWNLSAAIDS
YRKEKTEETRNLALIEEQATYRNAIHDFIGRTDNLTDANKRHAETIKGLFKAELFNGKVLKQLGTVTTTEHENALLRSFD
KFTTYFSGFYENRKNVFS AEDISTAI PHRIVQDNFPKF KENCHIFTRLITAVPSLREHFENVKKAIGIFVSTSI EEFVSFP
FYNQLLTQTQIDLYNQLLGGISREAGTEKIKGLNEVLNLAIQKNDETAHIIASLPHRFIPLFKQILSDRNTLSFILEEFKS
10 DEEVIQSFCKYKTLRLNENVLETAELFNELSIDLTHIFISHKKLETISSALCDHWDTLRNALYERRISELTGKITKSAK
EKVQSRSLKHEDINLQEII SAAGKELSEAFKQKTSEILSHAHAAALDQPLPTTLKKQEEKEILKSQLDSSLGLYHLLDWFVAVD
ESNEVDPEFSARLTGIKLEMEPSLSFYNKARNYATKKPYSVEKFKLNFQMPTLASGWDVNKEKNNGAILFVKNGLYLGLIM
PKQKGRYKALSFEPTKTSSEGFDMYYDYFPDAKMI PKCSTQLKAVTAHFQTHHTPIILLSNFI EPLEITKEIYDLNNPE
KEPKKFQTAYAKKTGDQKGYREALCKWIDFTRDFLSKYTKTTSIDLSSLRPSSQYKDLGEYYAELNPLLYHISFQRIAEKE
15 IMDAVETGKLYLFQIYNKDFAKGHHGKPNLHTLYWTGLFSPENLAKTSIKLNGQAE LFYRPKSRMKRMAHRLGEKMLNKKL
KDQKTPIPD TLYQELYDYVNHRLSHDLSDEARALLPNVITKEVSHEI IKDRRFTSDKFFFHVPITLNYQAANS PSKFNQRV
NAYLKEHPETPIIGIDRGERNLIYITVIDSTGKILEQ RSLNTIQQFDYQKKLDNREKERVAAARQAWSVVGTIKDLKQGYLS
QVIHEIVDLMIHYQAVVLENLNF GFKSKRTGIAEKAVYQQFEKMLIDKLNCLVLKDYPAEKVGGLNPNYQLTDQFTSFAK
MGTQSGFLFYVPAPYTSKIDPLTG FVDPFVWKT IKNHESRKHFLEGFDFLHYDVKTGDFILHFKMNRNLSFQ RGLPGFMFA
20 WDIVFEKNETQF DAKGTPFIAGKRIVPVIENHRFTGRYRDLYPANELIALLEEKGIVFRDGSN ILPKLEND DSHAITDMV
ALIRSVLQMRNSNAATGEDYINSPVRDLNGVC FDSRFQNPWPMDADANGAYHIALKGQLLLNHLKESKDLKLQNGISNQD
WLAYIQELRN
(SEQ ID NO: 25).

[0108] In embodiments, the RNA-guided DNA endonuclease enzyme is a Class II CRISPR endonuclease. A Class II CRISPR endonuclease can be identified via a method known in the art.

25 Type II CRISPR–Cas systems are defined by the presence of a single subunit crRNA–effector module. The Cpf1 enzyme belongs to a putative type V CRISPR–Cas system. Both type II and type V systems are included in Class II of the CRISPR–Cas system. Type II system are specific to bacteria whereas type V system is present in at least one archeon, see Makarova Nat Rev Microbiol 2015.

30 **[0109]** In embodiments, the RNA-guided DNA endonuclease enzyme has no nuclear localization sequence.

[0110] In embodiments, the Cas9 is a Cas9 variant, wherein the Cas9 variant has one or more mutations that increase its binding specificity to PAM compared to wild type Cas9. Without wishing to be bound by theory, it is believed that the target sequence should be associated with a PAM (protospacer adjacent motif); that is, a short sequence recognized by the CRISPR complex. The precise sequence and length requirements for the PAM differ depending on the CRISPR enzyme used, but PAMs are typically 2-5 base pair sequences adjacent the protospacer (that is, the target

sequence). Examples of PAM sequences are given in Table 2, and the skilled person will be able to identify further PAM sequences for use with a given CRISPR enzyme. The binding specificity of Cas9 or Cas9 variants to PAM can be determined by any method known in the art, for example, methods described in the references in Table 2 above.

- 5 **[0111]** In embodiments, the RNA-guided DNA endonuclease enzyme is *Staph aureus* Cas9, containing the amino acid sequence of SEQ ID NO: 26 or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 26 (or appropriate portion thereof).

10 KRNYILGLDIGITSVGYGIIDYETRDVIDAGVRLFKEANVENNEGRRSKRGARRLKRRRRHRIQRVKLLFDYNLLTDHSE
 LSGINPYEARVKGLSQKLSEEEFSAALLHLAKRRGVHNVNEVEEDTGNELSTREQISRNSKALEEKYVAELQLERLKKDGE
 VRGSINRFKTSYVKEAKQLLKVKQAYHQLDQSFIDTYIDLLETRRTYYEGPGEGSPFGWKDIKEWYEMLMGHCTYFPEEL
 RSVKYAYNADLYNALNDLNNLVITRDENEKLEYEKFQIIENVFKQKKKPTLKQIAKEILVNEEDIKGYRVTSTGKPEFTN
 LKVYHDIKDITARKEIIEAELLDQIAKILTIYQSSEDIQEELTNLNSLTQEEIEQISNLKGYTGTHNLSLKAINLILDE
 LWHTNDNQIAIFNRLKLVKKVLDLSQQKEIPTTLVDDFILSPVVKRSFIQSIKVINAIKKYGLPNDIIIEELAREKNSKDA
 15 QKMINEMQKRNRQTNERIEEIIIRTTGKENAKYLIKIKLHDMQEGKCLYSLEAIPLEDLLNNPFNYEVDHIIIPRSVSFDNS
 FNNKVLVKQEENSKKGNRTPFQYLSSSDSKISYETFKKHILNLAKGKGRISKTKEYLLEERDINRFSVQKDFINRNLVDT
 RYATRGLMNLRSYFRVNNLDVKVKSINGGFTSFLRRKWKFKKERNKGYKHHAEADALIIANADFIKWKKLDKAKKVMEN
 QMFEEKQAESMPEIETEQEYKEIFITPHQIKHIKDFKDYKYSHRVDKKPNRELINDTLYSTRKDDKGNTLIVNNLNGLYDK
 DNDKLLKLINKSPEKLLMYHHDPTQYQKLKLIMEQYGDEKNPLYKYEEETGNYLTYSKKDNGPVIKKIKYYGNKLNALHD
 20 ITDDYPNSRNKVKLSLKPYRFDVYLDNGVYKFVTVKNLDVIKKENYYEVNSKCYEEAKKLKKISNQAEFIASFYNNDLIK
 INGELYRVIGVNNDLLNRIEVNMIDITYREYLENMNDKRPPRIIKTIASKTQSIKKYSTDILGNLYEVKSKKHPQIIKKG
 (SEQ ID NO: 26)

- [0112]** In embodiments, the RNA-guided DNA endonuclease enzyme is *SpCas9* VQR mutant, containing the amino acid sequence of SEQ ID NO: 27 or an amino acid sequence having at least
 25 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 27 (or appropriate portion thereof).

30 DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL
 QEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLRILIYALAHMIK
 FRGHFLIEGDLNPDNSVDKLFQVLQVTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
 NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGAS
 AQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY

FKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKYAHLFDDKVMK
 QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
 PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKL
 YLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLI
 5 TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KvreINNYHHAHDAYLNAVVGtalIKKYPKLESEFVYGDYKVYDVRKMIakSEQEIGKATAKYFFYSNIMNFFKTEITLAN
 GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**
 SPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
 AGELQKGNELALPSKYVNFLYLASHYEKLKGSPEdNEQKQLFVEQHKKHYLDEIIEQISEFSKRVIADANLDKVLsAYNKH
 10 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK**QYR**STKEVL DATLIHQsITGLYETRIDLSQLGGD (SEQ ID
 NO: 27) [Bold and underlined residues indicate the mutations D1135V, R1335Q and
 T1337R]

[0113] In embodiments, the RNA-guided DNA endonuclease enzyme is SpCas9 VRER mutant,
 containing the amino acid sequence of SEQ ID NO: 28 or an amino acid sequence having at least
 15 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 28 (or appropriate
 portion thereof).

DKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL
 QEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 20 SLGLTPNFKSNFDLAEDAKQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
 NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGPLARGNSRFAWMTRKSEETITPWNFEVVDKGAS
 AQSFIERMTNFDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY
 FKKIECFDSVEISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIERLKYAHLFDDKVMK
 25 QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFQMQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
 PAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKL
 YLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKMKNYWRQLLNAKLI
 TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSCLVSDFRKDFQFY
 KvreINNYHHAHDAYLNAVVGtalIKKYPKLESEFVYGDYKVYDVRKMIakSEQEIGKATAKYFFYSNIMNFFKTEITLAN
 30 GEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGF**V**
 SPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSEFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGRKRMLAS
ARELQKGNELALPSKYVNFLYLASHYEKLKGSPEdNEQKQLFVEQHKKHYLDEIIEQISEFSKRVIADANLDKVLsAYNKH
 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK**EYR**STKEVL DATLIHQsITGLYETRIDLSQLGGD (SEQ ID
 NO: 28) [Bold and underlined residues indicate the mutations D1135V, G1218R,
 35 R1335E and T1337R]

[0114] In embodiments, the RNA-guided DNA endonuclease enzyme is SpCas9 D1135E mutant, containing the amino acid sequence of SEQ ID NO: 29 or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 29 (or appropriate portion thereof).

5 DKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL
 QEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
 10 NGSIHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGAS
 AQSFIERM TNF DKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY
 FKKIECFDSVEISGVEDRFNASLGT YHDLKKI IKDKDFLDNEENEDILEDIVLTLT LFEDREMI EERLKYAHLFDDKVMK
 QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKS DGFANRNF MQLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
 PAIKKGILQTVKVDELVKVMGRHKPENIVIE MARENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKL
 15 YLYYLQNGRDMYVDQELDINRLSDYD VDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKNYWRQLLNAKLI
 TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFY
 KVREINNYHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIKSEQEI GKATAKYFFYSNIMNFFKTEITLAN
 GEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFE
 SPTVAYSVLVVAKEGKSKKLKSVKELLGITIMERS SF EKNPIDFLEAKGYKEVKKDLIIKLPKYSLFEL ENGRKRMLAS
 20 AGELQKGNELALPSKYVNFLYLASHYEKLGSPEDNEQKQLFVEQHKHYLDEIIIEQISEFSKR VILADANLDKVL SAYNKH
 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVLDATLIHQ SITGLYETRIDLSQLGGD (SEQ ID
 NO: 29) [Bold and underlined residue indicates the mutation D1135E]

[0115] In embodiments, the RNA-guided DNA endonuclease enzyme is eSpCas9 1.1 mutant, containing the amino acid sequence of SEQ ID NO: 30 or an amino acid sequence having at least
 25 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 30 (or appropriate portion thereof).

DKKYSIGLDIGTNSVGWAVITDEYKVPSSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL
 QEIFSNEMAKVDDSFHRLSEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 30 SLGLTPNFKSNFDLAEDAKLQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRTFD
 NGSIHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFEVVVDKGAS
 AQSFIERM TNF DKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY
 FKKIECFDSVEISGVEDRFNASLGT YHDLKKI IKDKDFLDNEENEDILEDIVLTLT LFEDREMI EERLKYAHLFDDKVMK

QLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMLIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
 PAIKKGILQTVKVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EG IKELGSQILKEHPVENTQLQNEKL
 YLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFL**A**DDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLI
 TQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFY
 5 KVINNYHHHAHDAYLNAVVG TALIKKYP**A**LESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
 GEIRK**A**PLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD
 SPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYS LFELENGRKRMLAS
 AGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH
 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD (SEQ ID
 10 NO: 30) [Bold and underlined residues indicate the mutations K848A/K1003A/R1060A]

[0116] In embodiments, the RNA-guided DNA endonuclease enzyme is SpCas9 HF1 mutant, containing the amino acid sequence of SEQ ID NO: 31 or an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 31 (or appropriate portion thereof).

DKKYSIGLDIGTNSVGWAVITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL
 QEIFSNEMAKVDDSFHRL EESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLR LIY LALAHMIK
 FRGHFLIEGDLNPDNSDVKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSKSRLENLIAQLPGEKKNGLFGNLIAL
 SLGLTPNFKSNFDLAEDAKQLSKD TYDDDLNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRY
 DEHQDLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKF IKPILEKMDGTEELLVKLNREDLLRKQRTFD
 20 NGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYYVGPLARGNSRFAWMTRKSEETITPWNFE EVVDKGAS
 AQSFIERMT**A**FDKNLPNEKVLPHKSLLEYFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDY
 FKKIECFDSVEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMI EERLKTYAHLFDDKVMK
 QLKRRRYTGWGA**A**LSRKLINGIRDKQSGKTILDFLKSDGFANRNFMA**A**LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGS
 PAIKKGILQTVKVDELVKVMGRHKPENIVIEMARENQTTQKGQKNSRERMKRIE EG IKELGSQILKEHPVENTQLQNEKL
 25 YLYYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKMKKNYWRQLLNAKLI
 TQRKFDNLTKAERGGLSELDKAGFIKRQLVETR**A**ITKHVAQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFY
 KVINNYHHHAHDAYLNAVVG TALIKKYPKLESEFVYGDYKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLAN
 GEIRKRPLIETNGETGEIVWDKGRDFATVRKVL SMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFD
 SPTVAYSVLVAKVEKGKSKKLKSVKELLGITIMERSSSF EKNPIDFLEAKGYKEVKKDLIIKLPKYS LFELENGRKRMLAS
 30 AGELQKGNELALPSKYVNFLYLASHYEKLKGS PEDNEQKQLFVEQHKHYLDEIIEQISEFSKRVILADANLDKVL SAYNKH
 RDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRKRYTSTKEVL DATLIHQ SITGLYETRIDLSQLGGD (SEQ ID
 NO: 31) [Bold and underlined residues indicate the mutations Q695A, Q926A, N497A
 and R661A]

[0117] In embodiments, the composition described herein includes a delivery vehicle (e.g., a
 35 nanoparticle, a lipid particle or a viral vector) and a protein that comprises a mitochondrial

localization amino acid sequence covalently attached to Cas9, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence of SEQ ID NO: 1 and the Cas9 protein includes an amino acid sequence of SEQ ID NO: 24.

- 5 **[0118]** In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to Cpf1, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and the
- 10 Cpf1protein includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 24.

- [0119]** In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to Cpf1, where the protein is bound to the delivery vehicle,
- 15 the mitochondrial localization amino acid sequence includes an amino acid sequence of SEQ ID NO: 1 and the Cpf1protein includes an amino acid sequence of SEQ ID NO: 25.

- [0120]** In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the
- 20 delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 25.

- [0121]** In embodiments, the composition described herein includes a delivery vehicle (e.g., a
- 25 nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence of SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence of SEQ ID NO: 26.

[0122] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 26.

[0123] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence of SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence of SEQ ID NO: 27.

[0124] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 27.

[0125] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence of SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence of SEQ ID NO: 28.

[0126] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the

delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 28.

5 **[0127]** In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence of SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence of SEQ ID
10 NO: 29.

[0128] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid
15 sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 29.

[0129] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization
20 amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence of SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence of SEQ ID NO: 30.

[0130] In embodiments, the composition described herein includes a delivery vehicle (e.g., a
25 nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID

NO: 1 and the Cas9 variant protein includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 30.

[0131] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence of SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence of SEQ ID NO: 31.

[0132] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein that contains a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the protein is bound to the delivery vehicle, the mitochondrial localization amino acid sequence includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and the Cas9 variant protein includes an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 31.

[0133] In another aspect, there is provided a composition including a delivery vehicle and a nucleic acid encoding a protein as disclosed herein. In embodiments, the composition includes a delivery vehicle and a nucleic acid encoding a protein according to any aspect or embodiment disclosed above.

[0134] In embodiments, the composition includes a delivery vehicle and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence (e.g., SEQ ID NO:1) covalently attached to an RNA-guided DNA endonuclease enzyme (e.g., Cas9, Cpf1, or a Class II CRISPR endonuclease), where the nucleic acid is bound to the delivery vehicle.

[0135] In embodiments, a nucleic acid encoding a mitochondrial localization amino acid sequence includes the nucleic acid sequence of SEQ ID NO: 2 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2.

[0136] In embodiments, a nucleic acid encoding Cas9 includes the nucleic acid sequence of SEQ ID NO: 33 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 32.

gacaagaagt acagcatcgg cctggacatc ggcaccaact ctgtgggctg
 ggccgtgatc accgacgagt acaaggtgcc cagcaagaaa ttcaaggtgc tgggcaaacac
 cgaccggcac agcatcaaga agaacctgat cggagccctg ctgttcgaca gcggcgaaac
 agccgaggcc acccggtga agagaaccgc cagaagaaga tacaccagac ggaagaaccg
 5 gatctgctat ctgcaagaga tcttcagcaa cgagatggcc aaggtggacg acagcttctt
 ccacagactg gaagagtcct tcctggtgga agaggataag aagcacgagc ggcaccccat
 cttcggcaac atcgtggacg aggtggccta ccacgagaag taccaccacca tctaccacct
 gaaaaagaaa ctggtggaca gcaccgacaa ggccgacctg cggtgatct atctggccct
 10 ggcccacatg atcaagttcc ggggccactt cctgatcgag ggcgacctga accccgacaa
 cagcgacgtg gacaagctgt tcatccagct ggtgcagacc tacaaccagc tgttcgagga
 aaaccccatc aacgccagcg gcgtggacgc caaggccatc ctgtctgcca gactgagcaa
 gagcagacgg ctggaatac tgatcgccca gctgcccggc gagaagaaga atggcctgtt
 cggaacactg attgcctga gcctgggcct gacccccaac ttcaagagca acttcgacct
 ggccgaggat gccaaactgc agctgagcaa ggacacctac gacgacgacc tggacaacct
 15 gctggcccag atcggcgacc agtacgccga cctgtttctg gccgccaaga acctgtccga
 cgccatcctg ctgagcgaca tcctgagagt gaacaccgag atcaccaagg ccccccctgag
 cgctctatg atcaagagat acgacgagca ccaccaggac ctgacctgc tgaagctct
 cgtgcccag cagctgcctg agaagtacaa agagattttc ttcgaccaga gcaagaacgg
 ctacgccggc tacattgacg gcggagccag ccaggaagag ttctacaagt tcatcaagcc
 20 catcctggaa aagatggacg gcaccgagga actgctcgtg aagctgaaca gagaggacct
 gctgcggaag cagcggacct tcgacaacgg cagcatcccc caccagatcc acctgggaga
 gctgcacgcc attctgcggc ggcaggaaga tttttacca ttctgaagg acaaccggga
 aaagatcgag aagatcctga ccttcgcgcat cccctactac gtgggcccctc tggccagggg
 aaacagcaga ttgcctgga tgaccagaaa gagcgaggaa accatcacc cctggaactt
 25 cgaggaagtg gtggacaagg gcgcttcgc ccagagcttc atcgagcgga tgaccaactt
 cgataagaac ctgcccacg agaaggtgct gcccaagcac agcctgctgt acgagtactt
 caccgtgtat aacgagctga ccaaagtga atacgtgacc gagggaatga gaaagcccgc
 cttcctgagc ggcgagcaga aaaaggccat cgtggacctg ctgttcaaga ccaaccggaa
 agtgaccgtg aagcagctga aagaggacta cttcaagaaa atcgagtgt tgcactccgt
 30 ggaaatctcc ggcgtggaag atcggttcaa cgctccctg ggcacatacc acgatctgct
 gaaaattatc aagggaacgg acttccctgga caatgaggaa aacgaggaca ttctggaaga
 tatcgtgctg accctgacac tgtttgagga cagagagatg atcgaggaaac ggctgaaaac
 ctatgccac ctgttcgacg acaaagtgat gaagcagctg aagcggcgga gatacacccg
 ctggggcagg ctgagccgga agctgatcaa cggcatccgg gacaagcagt ccggcaagac
 35 aatcctggat ttctgaagt ccgacggctt cgccaacaga aacttcatgc agctgatcca
 cgacgacagc ctgaccttta aagaggacat ccagaaagcc caggtgtccg gccaggcgga
 tagcctgcac gagcacattg ccaatctggc cggcagcccc gccattaaga agggcatcct
 gcagacagtg aaggtggtgg acgagctcgt gaaagtgat ggccggcaca agcccgagaa
 catcgtgatc gaaatggcca gagagaacca gaccaccag aagggacaga agaacagccg
 40 cgagagaatg aagcggatcg aagagggcat caaagagctg ggcagccaga tcctgaaaga
 acaccccgtg gaaaacaccc agctgcagaa cgagaagctg tacctgtact acctgcagaa
 tgggcgggat atgtacgtgg accaggaact ggacatcaac cggctgtccg actacgatgt
 ggaccatata gtgcctcaga gctttctgaa ggacgactcc atcgacaaca aggtgctgac
 45 cagaagcgac aagaaccggg gcaagagcga caacgtgcc tccgaagagg tcgtgaagaa
 gatgaagaac tactggcggc agctgctgaa cgccaagctg attaccaga gaaagttcga
 caatctgacc aaggccgaga gaggcggcct gagcgaactg gataaggccg gcttcatcaa
 gagacactg gtggaaccc ggcatcac aaagcacgtg gcacagatcc tggactcccg
 gatgaacact aagtacgacg agaatgacaa gctgatccgg gaagtgaag tgatcacct
 gaagtccaag ctggtgtccg atttccggaa ggatttccag ttttacaag tgcgcgagat
 50 caacaactac caccacgccc acgacgccta cctgaacgcc gtcgtgggaa ccgcctgat
 caaaaagtac cctaagctgg aaagcgagtt cgtgtacggc gactacaagg tgtacgacgt
 gcggaagatg atcgccaaga gcgagcagga aatcggcaag gctaccgcca agtacttctt
 ctacagcaac atcatgaact ttttcaagac cgagattacc ctggccaacg gcgagatccg
 gaagcggcct ctgatcgaga caaacggcga aaccggggag atcgtgtggg ataaggccg
 55 ggattttgcc accgtgcgga aagtgtgtgag catgcccac gtgaatatcg tgaaaaagac
 cgaggtgcag acaggcggct tcagcaaaga gtctatcctg cccaagagga acagcgataa

gctgatcgcc agaaagaagg actgggaccc taagaagtag ggcgggttcg acagccccac
 cgtggcctat tctgtgctgg tgggtggccaa agtggaaaag ggcaagtcca agaaactgaa
 gagtgtgaaa gagctgctgg ggatcaccat catggaaaaga agcagcttcg agaagaatcc
 catcgacttt ctggaagcca agggctacaa agaagtgaaa aaggacctga tcatcaagct
 5 gcctaagtac tccctgttcg agctggaaaa cgcccggaag agaatgctgg cctctgccgg
 cgaactgcag aagggaacg aactggccct gccctccaaa tatgtgaact tcctgtacct
 ggccagccac tatgagaagc tgaagggctc ccccaggat aatgagcaga aacagctgtt
 tgtggaacag cacaagcact acctggacga gatcatcgag cagatcagcg agttctccaa
 gagagtgatc ctggccgacg ctaatctgga caaagtgtg tccgcctaca acaagcaccg
 10 ggataagccc atcagagagc aggcgagaa tatcatccac ctgtttaccc tgaccaatct
 gggagccct gccgccttca agtactttga caccaccatc gaccggaaga ggtacaccag
 caccaaagag gtgctggacg ccacctgat ccaccagagc atcaccggcc tgtacgagac
 acgcatcgac ctgtctcagc tgggagcgca c (SEQ ID NO: 32)

- 15 **[0137]** In embodiments, a nucleic acid encoding Cpf1 includes the nucleic acid sequence of SEQ ID NO: 33 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 33.

ACACAGTTCGAGGGCTTTACCAACCTGTATCAGGTGAGCAAGACACTGCGGTTTGAGCTGATCCACAGGGCAAGACCCTG
 AAGCACATCCAGGAGCAGGGCTTCATCGAGGAGGACAAGGCCCGCAATGATCACTACAAGGAGCTGAAGCCCATCATCGAT
 20 CGGATCTACAAGACCTATGCCGACAGTGCCTGCAGCTGGTGCAGCTGGATTGGGAGAACCTGAGCGCCGCCATCGACTCC
 TATAGAAAGGAGAAAACCGAGGAGACAAGGAACGCCCTGATCGAGGAGCAGGCCACATATCGCAATGCCATCCACGACTAC
 TTCATCGGCCGACAGACAACCTGACCGATGCCATCAATAAGAGACACGCCGAGATCTACAAGGGCCTGTTCAAGGCCGAG
 CTGTTTAAATGGCAAGGTGCTGAAGCAGCTGGGCACCGTGACCACAACCGAGCACGAGAACGCCCTGCTGCGGAGCTTCGAC
 AAGTTTACAACCTACTTCTCCGGCTTTTATGAGAACAGGAAGAAGCTGTTTCAGCGCCGAGGATATCAGCACAGCCATCCCA
 25 CACCGCATCGTGCAGGACAACCTCCCAAGTTTAAAGGAGAATTGTACATCTTCACACGCCTGATCACCGCCGTGCCAGC
 CTGCGGGAGCACTTTGAGAACGTGAAGAAGGCCATCGGCATCTTCGTGAGCACCTCCATCGAGGAGGTGTTTTCTTCCCT
 TTTTATAACCAGCTGCTGACACAGACCCAGATCGACCTGTATAACCAGCTGCTGGGAGGAATCTCTCGGGAGGCAGGCACC
 GAGAAGATCAAGGGCCTGAACGAGGTGCTGAATCTGGCCATCCAGAAGAATGATGAGACAGCCACATCATCGCTCCCTG
 CCACACAGATTCATCCCCCTGTTTAAGCAGATCCTGTCCGATAGGAACACCCTGTCTTTCATCCTGGAGGAGTTTAAGAGC
 30 GACGAGGAAGTGATCCAGTCTTCTGCAAGTACAAGACACTGCTGAGAAACGAGAACGTGCTGGAGACAGCCGAGGCCCTG
 TTTAACGAGCTGAACAGCATCGACCTGACACACATCTTCATCAGCCACAAGAAGCTGGAGACAATCAGCAGCGCCCTGTGC
 GACCACTGGGATACACTGAGGAATGCCCTGTATGAGCGGAGAATCTCCGAGCTGACAGGCAAGATCACCAAGTCTGCCAAG
 GAGAAGGTGCAGCGCAGCCTGAAGCACGAGGATATCAACCTGCAGGAGATCATCTCTGCCGAGGCAAGGAGCTGAGCGAG
 GCCTTCAAGCAGAAAACCAGCGAGATCCTGTCCACGCACACGCCGCCCTGGATCAGCCACTGCCTACAACCCTGAAGAAG
 35 CAGGAGGAGAAGGAGATCCTGAAGTCTCAGCTGGACAGCCTGCTGGGCCTGTACCACCTGCTGGACTGGTTTGCCGTGGAT
 GAGTCCAACGAGGTGGACCCCGAGTTCTTGCCCGGCTGACCGGCATCAAGCTGGAGATGGAGCCTTCTCTGAGCTTCTAC
 AACAAGGCCAGAAATTATGCCACCAAGAAGCCCTACTCCGTGGAGAAGTTCAAGCTGAACTTTCAGATGCCTACACTGGCC
 TCTGGCTGGGACGTGAATAAGGAGAAGAACAATGGCGCCATCCTGTTTGTGAAGAACGGCCTGTACTATCTGGGCATCATG
 CCAAAGCAGAAGGGCAGGTATAAGGCCCTGAGCTTCGAGCCACAGAGAAAACCAGCGAGGGCTTTGATAAGATGTACTAT
 40 GACTACTTCCCTGATGCCGCCAAGATGATCCCAAAGTGACGACCCAGCTGAAGGCCGTGACAGCCCACTTTCAGACCCAC
 ACAACCCCCATCCTGCTGTCCAACAATTTTCATCGAGCCTCTGGAGATCACAAAGGAGATCTACGACCTGAACAATCCTGAG

AAGGAGCCAAAGAAGTTTCAGACAGCCTACGCCAAGAAAACCGGCGACCAGAAGGGCTACAGAGAGGCCCTGTGCAAGTGG
 ATCGACTTCACAAGGGATTTTCTGTCCAAGTATACCAAGACAACCTCTATCGATCTGTCTAGCCTGCGGCCATCCTCTCAG
 TATAAGGACCTGGGCGAGTACTATGCCGAGCTGAATCCCCTGCTGTACCACATCAGCTTCCAGAGAATCGCCGAGAAGGAG
 ATCATGGATGCCGTGGAGACAGGCAAGCTGTACCTGTTCCAGATCTATAACAAGGACTTTGCCAAGGGCCACCACGGCAAG
 5 CCTAATCTGCACACACTGTATTGGACCGGCCTGTTTTCTCCAGAGAACCTGGCCAAGACAAGCATCAAGCTGAATGGCCAG
 GCCGAGCTGTTCTACCGCCCTAAGTCCAGGATGAAGAGGATGGCACACCGGCTGGGAGAGAAGATGCTGAACAAGAAGCTG
 AAGGATCAGAAAACCCCAATCCCCGACACCCTGTACCAGGAGCTGTACGACTATGTGAATCACAGACTGTCCCACGACCTG
 TCTGATGAGGCCAGGGCCCTGCTGCCCAACGTGATCACCAGGAGGTGTCTCACGAGATCATCAAGGATAGGCGCTTTACC
 AGCGACAAGTTCTTTTTCCACGTGCCTATCACACTGAAGTATCAGGCCGCCAATTCCCCATCTAAGTTCAACCAGAGGGTG
 10 AATGCCTACCTGAAGGAGCACCCCGAGACACCTATCATCGGCATCGATCGGGGCGAGAGAAACCTGATCTATATCACAGTG
 ATCGACTCCACCGCAAGATCCTGGAGCAGCGGAGCCTGAACACCATCCAGCAGTTTGATTACCAGAAGAAGCTGGACAAC
 AGGGAGAAGGAGAGGGTGGCAGCAAGGCAGGCCTGGTCTGTGGTGGGCACAATCAAGGATCTGAAGCAGGGCTATCTGAGC
 CAGGTCATCCACGAGATCGTGGACCTGATGATCCACTACCAGGCCGTGGTGGTCTGGAGAACCTGAATTTTCGGCTTTAAG
 AGCAAGAGGACCGGCATCGCCGAGAAGGCCGTGTACCAGCAGTTCGAGAAGATGCTGATCGATAAGCTGAATTGCCTGGTG
 15 CTGAAGGACTATCCAGCAGAGAAAGTGGGAGGCGTGTGAACCCATACCAGCTGACAGACCAGTTCACCTCCTTTGCCAAG
 ATGGGCACCCAGTCTGGCTTCCTGTTTTACGTGCCTGCCCCATATACATCTAAGATCGATCCCCTGACCGGCTTCGTGGAC
 CCCTTCGTGTGAAAACCATCAAGAATCACGAGAGCCGCAAGCACTTCCTGGAGGGCTTCGACTTTCTGCACTACGACGTG
 AAAACCGGCGACTTCATCCTGCACCTTAAGATGAACAGAAATCTGTCCTTCCAGAGGGGCTGCCCGGCTTTATGCCTGCA
 TGGGATATCGTGTTTCGAGAAGAACGAGACACAGTTTGACGCCAAGGGCACCCCTTTCATCGCCGGCAAGAGAATCGTGCCA
 20 GTGATCGAGAATCACAGATTACCGGCAGATACCGGGACCTGTATCCTGCCAACGAGCTGATCGCCCTGCTGGAGGAGAAG
 GGCATCGTGTTTCAGGGATGGCTCCAACATCCTGCCAAAGCTGCTGGAGAATGACGATTCTCACGCCATCGACACCATGGTG
 GCCCTGATCCGCAGCGTGTGTCAGATGCGGAACCTCCAATGCCGCCACAGGCGAGGACTATATCAACAGCCCCGTGCGCGAT
 CTGAATGGCGTGTGCTTCGACTCCCGGTTTCAGAACCCAGAGTGGCCCATGGACGCCGATGCCAATGGCGCCTACCACATC
 GCCCTGAAGGGCCAGCTGCTGCTGAATCACCTGAAGGAGAGCAAGGATCTGAAGCTGCAGAACGGCATCTCCAATCAGGAC
 25 TGGCTGGCCTACATCCAGGAGCTGCGCAACAAAAGGCCGGCGGCCACGAAAAAGGCCGGCCAGGCAAAAAAGAAAAAG
 (SEQ ID NO: 33)

[0138] In embodiments, a nucleic acid encoding a Cas9 variant (SaCas9) includes the nucleic acid sequence of SEQ ID NO: 34 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 34.

aagcggaaactacatcctgggcctggacatcggcatcaccagcgtgggctacggcatcatcgactacgagacacgggacgtg
 atcgatgccggcgctgcggctgttcaaagaggccaacgtggaaaacaacgagggcgaggcgagcaagagagggcgccagaagg
 ctgaagcggcgaggcgcatagaatccagagagtggaagaagctgctgttcgactacaacctgctgaccgaccacagcgag
 ctgagcggcatcaaccctacgaggccagagtgaagggcctgagccagaagctgagcgaggaagagttctctgccgccctg
 ctgcacctggccaagagaagaggcgtgcacaacgtgaacgaggtggaagaggacaccggcaacgagctgtccaccagagag
 35 cagatcagccggaacagcaaggccctggaagagaaatacgtggccgaactgcagctggaacggctgaagaaagacggcgaa
 gtgcggggcagcatcaacagattcaagaccagcgactacgtgaagaagccaacagctgctgaaggtgcagaaggcctac

caccagctggaccagagcttcatcgacacctacatcgacctgctggaaacccggcggaacctactatgagggacctggcgag
 ggcagccccttcggctggaaggacatcaaagaatggtacgagatgctgatgggccactgcacctacttccccgaggaactg
 cggagcgtgaagtacgcctacaacgccgacctgtacaacgccctgaacgacctgaacaatctcgtgatcaccagggacgag
 aacgagaagctggaatattacgagaagttccagatcatcgagaacgtgttcaagcagaagaagaagcccaccctgaagcag
 5 atcgccaaagaaatcctcgtgaacgaagaggatattaagggctacagagtgaccagcaccggcaagcccaggttcaccaac
 ctgaaggtgtaccacgacatcaaggacattaccgcccggaaagagattattgagaacgccgagctgctggatcagattgcc
 aagatcctgaccatctaccagagcagcgaggacatccaggaagaactgaccaatctgaactccgagctgaccacaggaagag
 atcgagcagatctctaattctgaagggctataccggcaccacacaacctgagcctgaaggccatcaacctgatcctggacgag
 ctgtggcacaccaacgacaaccagatcgctatcttcaaccggctgaagctggtgcccagaaggtggacctgtcccagcag
 10 aaagagatccccaccaccctggtggacgacttcatcctgagccccgctcgtgaagagaagcttcatccagagcatcaaagtg
 atcaacgccatcatcaagaagtacggcctgcccacgacatcattatcgagctggcccgcgagaagaactccaaggacgcc
 cagaaaatgatcaacgagatgcagaagcggaaaccggcagaccaacgagcggatcgaggaaatcatccggaccaccggcaaa
 gagaacgccaaagtacctgatcgagaagatcaagctgcacgacatgcaggaaggcaagtgcctgtacagcctggaagccatc
 cctctggaagatctgctgaacaaccccttcaactatgaggtggaccacatcatccccagaagcgtgtccttcgacaacagc
 15 ttcaacaacaaggtgctcgtgaagcaggaagaaaacagcaagaagggcaaccggacccccattccagtagctgagcagcagc
 gacagcaagatcagctacgaaaccttcaagaagcacatcctgaatctggccaagggcaagggcagaatcagcaagaccaag
 aaagagtatctgctggaagaacgggacatcaacaggttctccgtgcagaaagacttcatcaaccggaaacctggtggatacc
 agatacgccaccagaggcctgatgaacctgctgcggagctacttcagagtgaacaacctggacgtgaaagtgaagtccatc
 aatggcggcttcaccagctttctgcggcggaagtggaaagtttaagaaagagcgggaacaaggggtacaagcaccacgcccag
 20 gacgcctgatcattgccaacgccgatttcatcttcaagagtggaaagaaactggacaaggccaaaaaagtgatggaaaac
 cagatgttcgaggaaaagcaggccgagagcatgcccagatcgaaaccgagcaggagtacaaagagatcttcatcaccccc
 caccagatcaagcacattaaggacttcaaggactacaagtacagccaccgggtggacaagaagcctaatagagagctgatt
 aacgacacctgtactccaccgggaaggacgacaagggcaacacctgatcgtgaacaatctgaacggcctgtacgacaag
 gacaatgacaagctgaaaaagctgatcaacaagagccccgaaaagctgctgatgtaccaccacgacccccagacctaccag
 25 aaactgaagctgattatggaacagtagcggcgacgagaagaatccctgtacaagtactacgaggaaaccgggaactacctg
 accaagtactccaaaaaggacaacggccccgtgatcaagaagattaagtattacggcaacaaactgaacgcccattctggac
 atcaccgacgactacccaacagcagaaacaaggtcgtgaagctgtccctgaagccctacagattcgacgtgtacctggac
 aatggcgtgtacaagttcgtgaccgtgaagaatctggatgtgatcaaaaaagaaaactactacgaagtgaatagcaagtgc
 tatgaggaagctaagaagctgaagaagatcagcaaccaggccgagtttatcgccctccttctacaacaacgatctgatcaag
 30 atcaacggcgagctgtatagagtgatcggcgtgaacaacgacctgctgaaccggatcgaaagtgaacatgatcgacatcacc
 taccgagtagctggaacacatgaacgacaagagggccccccaggatcattaagacaatcgccccaagaccagagcatt
 aagaagtacagcacagacattctgggcaacctgtatgaagtgaatctaagaagcaccctcagatcatcaaaaagggc
 (SEQ ID NO: 34)

[0139] In embodiments, a nucleic acid encoding a Cas9 variant (SpCas9 VQR) includes the
 35 nucleic acid sequence of SEQ ID NO: 35 or a nucleic acid having at least 80%, 85%, 90%, 95%,
 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 35.

gataaaaagtattctatttggttagacatcggcactaattccgttggtggctgtcataaccgatgaatacaaaagtacct
tcaaagaaatttaaggtggtgggaacacagaccgtcattcgattaaaaagaatcttatcgggtgccctcctattcgatagt
ggcgaaacggcagaggcgactcgctgaaacgaaccgctcggagaaggtatacacgtcgcaagaaccgaatatgttactta
caagaaatTTTTtagcaatgagatggccaaagttgacgattctttctttcacggttggaagagtccttccttgtcgaagag
5 gacaagaacatgaacggcaccccatctttgaaacatagtagatgaggtggcatatcatgaaaagtacccaacgatttat
cacctcagaaaaaagctagttgactcaactgataaagcggacctgaggttaatctacttggtccttgcccatatgataaag
ttcctggtggcactttctcattgaggggtgatctaaatccggacaactcggatgtcgacaaactgttcatccagtttagtaciaa
acctataatcagttgtttgaagagaaccctataaatgcaagtggcgtggatgcaaggctattcttagcgccccgctctct
aaatcccgacggctagaaaacctgatcgacacaattaccggagagaagaaaaatgggttggttcggtaaccttatagcgctc
10 tctactaggcctgacaccaaattttaagtcgaacttcgacttagctgaagatgccaaattgcagcttagtaaggacacgtac
gatgacgatctcgacaatctactggcacaattggagatcagtatgaggacttatttttggtgcgcaaaaaccttagcgat
gcaatcctcctatctgacatactgagagtttaatactgagattaccaaggcgccgttatccgcttcaatgatcaaaagggtac
gatgaacatcaccaagacttgacacttctcaaggccctagtcggtcgcaactgcctgagaaatataaggaaatattcttt
gatcagtcgaaaaacgggtacgcaggttatattgacggcggagcgagtcagaggaattctacaagtttatcaaaccata
15 ttagagaagatggatgggacggaagagttgcttgtaaaactcaatcggaagatctactgcgaaagcagcggactttcgac
aacggtagcattccacatcaaatccacttaggcgaattgcatgctatacttagaaggcaggaggatttttatccgttcctc
aaagacaatcgtgaaaagattgagaaaatcctaacctttcgcataccttactatgtgggacccctggcccgagggaactct
cggttcgcatggatgacaagaaagtcggaagaaacgattactccatggaaattttgaggaagttgtcgataaagggtgcgtca
gctcaatcgttcatcgagaggatgaccaactttgacaagaatttaccgaacgaaaaagatttgccctaagcacagtttactt
20 tacgagtatttcacagtgtacaatgaactcacgaaagttaagtatgtcactgagggcatgcgtaaacccgcctttctaagc
ggagaacagaagaaagcaatagtagatctgttattcaagaccaaccgcaaagtgcagtttaagcaattgaaagaggactac
tttaagaaaattgaatgcttcgattctgtcgagatctccgggtagaagatcgatttaatgcgtcacttggtacgtatcat
gacctcctaagataattaaagataaggacttcctggataacgaagagaatgaagatatcttagaagatatagtgttgact
cttaccctctttgaagatcgggaaatgattgaggaaagactaaaaacatacgtcacctggttcgacgataagggttatgaaa
25 cagttaaagaggcgtcgctatacgggctggggacgattgtcgcggaacttatcaacgggataagagacaagcaaaagtgggt
aaaactattctcgattttctaaagagcgacggcttcgccaataggaactttatgcagctgatccatgatgactctttaacc
ttcaaagaggatatacaaaaggcacaggtttccggacaaggggactcattgcacgaacatattgcgaatcttgctggttcg
ccagccatcaaaaagggcatactccagacagtcaaaagtagtggatgagctagtttaaggatcatgggacgtcacaaccggaa
aacattgtaatcgagatggcacgcgaaaatcaaacgactcagaaggggcaaaaaaacagtcgagagcggatgaagagaata
30 gaagaggggtattaaagaactgggcagccagatcttaaaggagcatcctgtggaaaatacccaattgcagaacgagaaactt
tacctctattacctacaaaatggaaggacatgtatgttgatcaggaactggacataaacggtttatctgattacgacgtc
gatcacattgtaccccaatcctttttgaaggacgattcaatcgacaataaagtgttacacgctcggataagaaccgaggg
aaaagtgacaatgttccaagcgaggaagtcgtaagaaaaatgaagaactattggcggcagctcctaataatgcgaaactgata
acgcaaagaaagttcgataacttaactaaagctgagaggggtggcttgctgaacttgacaaggccggatttattaaacgt
35 cagctcgtggaacccgccaaatcacaagcatgttgacagatactagattcccgaatgaatacgaaatacgacgagaaac
gataagctgattcgggaagtcaaagtaatcactttaaagtcaaaattgggtgtcggaacttcagaaaggattttcaattctat
aaagttaggagataaataactaccacatgacgacgacgttatcttaatgccgtcgtagggaccgcactcattaagaaa
taccggaagctagaaagtgagtttggtgattacaaagtttatgacgtccgtaagatgatcgcgaaaagcgaaacag

gagataggcaaggctacagccaaatacttcttttattctaacattatgaatttctttaagacggaaatcactctggcaaac
 ggagagatacgcaaacgacctttaattgaaaccaatggggagacaggtgaaatcgtatgggataagggccgggacttcgcg
 acggtgagaaaagtgttccatgccccagtcaacatagtaaagaaaactgaggtgcagaccggagggttttcaaaggaa
 tcgattcttccaaaaaggaatagtataagctcatcgctcgtaaaaaggactgggacccgaaaaagtagcgttggtctctg
 5 agccctacagttgcctattctgtcctagtagtggtgcaaaagttgagaaggggaaaatccaagaaactgaagtcaagaa
 ttattgggataacgattatggagcgctcgtcttttgaaaagaaccccatcgacttccttgaggcgaaagggttacaaggaa
 gtaaaaaaggatctcataattaaactaccaaagtatagtctgtttgagttagaaaatggccgaaaaaggatgttggttagc
 gccggagagcttcaaaaggggaacgaactcgactacgtctaaatacgtgaatttctgtatttagcgtcccattacgag
 aagttgaaagggtcacctgaagataacgaacagaagcaactttttgttgagcagcacaacattatctcgacgaaatcata
 10 gagcaatttccgaattcagtaagagagtcctcctagctgatgccaatctggacaaagtattaagcgcatacaacaagcac
 agggataaaccctacgtgagcagggcgaaaatattatccatttgtttactcttaccaacctcggcgctccagccgcatc
 aagtattttgacacaacgatagatcgcaaacagtacagatctaccaaggaggtgctagacgcgacactgattcaccaatcc
 atcacgggattatatgaaactcggatagatttgtcacagcttgggggtgac (SEQ ID NO: 35)

[0140] In embodiments, a nucleic acid encoding a Cas9 variant (SpCas9 VRER) includes the
 15 nucleic acid sequence of SEQ ID NO: 36 or a nucleic acid having at least 80%, 85%, 90%, 95%,
 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 36.

gataaaaagtattctatttggttttagacatcggcactaattccgttggatgggctgtcataaccgatgaatacaaagtacct
 tcaaagaaatttaaggtgttggggaacacagaccgtcattcgattaaaaagaatcttatcgggtgccctcctattcgatagt
 ggcgaaacggcagaggcgactcgctgaaacgaaccgctcggagaagggtatacacgtcgcaagaaccgaatatgttactta
 20 caagaaattttttagcaatgagatggccaaagttgacgattctttctttcaccgtttggaagagtccttccttgtcgaagag
 gacaagaaacatgaacggcaccccatctttggaaacatagtagatgaggtggcatatcatgaaaagtacccaacgatttat
 cacctcagaaaaaagctagttgactcaactgataaagcggacctgaggttaatctacttggtccttgcccatatgataaag
 ttccgtgggcactttctcattgaggggtgatctaaatccggacaactcggatgtcgacaaactgttcacccagttagtacaa
 acctataatcagttgtttgaagagaacctataaatgcaagtggcgtggatgcaaggctattcttagcgccccgcctctct
 25 aaatcccgcagcgctagaaaacctgatcgacacaattaccggagagaagaaaaatgggttgttcggtaaccttatagcgctc
 tcactaggcctgacaccaaattttaagtcgaacttcgacttagctgaagatgcaaaattgcagcttagtaaggacacgtac
 gatgacgatctcgacaatctactggcacaattggagatcagtatgaggacttatttttggctgccaaaaaccttagcgat
 gcaatcctcctatctgacatactgagagtttaatactgagattaccaaggcgccgttatccgcttcaatgatcaaaaggtag
 gatgaacatcaccaagacttgacacttctcaaggccctagtcggtcagcaactgcctgagaaatataaggaaatattcttt
 30 gatcagtcgaaaaacgggtacgcaggttatattgacggcgagcgagtcagaggaattctacaagtttatcaaaccata
 ttagagaagatggatgggacggaagagttgcttgtaaaactcaatcggaagatctactgcgaaagcagcggactttcgac
 aacggttagcattccacatcaaattcacttaggcgaattgcatgctatacttagaaggcaggaggatttttatccgttcctc
 aaagacaatcgtgaaaagattgagaaaatcctaacctttcgcataccttactatgtgggacccctggcccgagggaactct
 cggttcgcatggatgacaagaaagtcgaagaaacgattactccatggaattttgaggaagttgtcgataaagggtgcgtca
 35 gctcaatcgttcacgagaggtgaccaactttgacaagaatttaccgaacgaaaaagtagtgcctaagcacagtttactt
 tacgagtatttcacagtgtagaatgaactcacgaaagttaagtagtgcactgagggcatgcgtaaacccgcctttctaaagc

ggagaacagaagaaagcaatagtagatctgttattcaagaccaaccgcaaagtgcaggttaagcaattgaaagaggactac
 ttttaagaaaattgaatgcttcgattctgtcgagatctccggggtagaagatcgatttaatgcgtcacttggtacgtatcat
 gacctcctaaagataattaaagataaggacttcctggataacgaagagaatgaagatatcttagaagatatagtgttgact
 cttaccctctttgaagatcgggaaatgattgaggaaagactaaaaacatacgcctcacctggttcgacgataagggttatgaaa
 5 cagttaaagaggcgctcgctatacgggctggggacgattgtcgcggaacttatcaacgggataagagacaagcaaagtgggt
 aaaactattctcgatttttctaagagcgacggcttcgccaataggaactttatgcagctgatccatgatgactctttaacc
 ttcaaagaggatatacaaaaggcacagggtttccggacaaggggactcattgcacgaacatattgcgaatcttgctggttcg
 ccagccatcaaaaagggcatactccagacagtc aaagtagtgatgagctagtttaagggtcatgggacgtc caaaaccggaa
 aacattgtaatcgagatggcacgcgaaaatcaaacgactcagaagggggcaaaaaaacagtcgagagcggatgaagagaata
 10 gaagaggggtattaaagaactgggcagccagatcttaaaggagcatcctgtggaaaatacccaattgcagaacgagaaaactt
 tacctctattacctacaaaatggaaggacatgtatggtgatcaggaactggacataaaaccgtttatctgattacgacgtc
 gatcacattgtaccccaatcctttttgaaggacgattcaatcgacaataaagtgccttacacgctcggataagaaccgaggg
 aaaagtgcacaaatgttccaagcgaggaagtcgtaaaagaaaatgaagaactattggcggcgagctcctaaatgcgaaactgata
 acgcaaagaaagttcgataacttaactaaagctgagaggggtggcttgtctgaacttgacaaggccggatttattaaacgt
 15 cagctcgtggaaacccgccaaatcacaagcatggttcacagatactagattcccgaatgaatacgaatacgcagcagaaac
 gataagctgattcgggaagtcaaagtaatcactttaaagtcaaaattgggtgctggacttcagaaaggattttcaattctat
 aaagttaggagataaataactaccacatgcgcacgacgcttatcttaatgccgtcgtagggaccgcactcattaagaaa
 taccgaagctagaaagttagttgtgtatggtgattacaaagtttatgacgtccgtaagatgatcgcgaaaagcgaaacag
 gagataggcaaggctacagccaaatacttcttttattctaacttatgaatttctttaagacggaaatcactctggcaaac
 20 ggagagatacgc aaacgacctttaattgaaaccaatggggagacaggtgaaatcgatgggataaggggccgggacttcgcg
 acgggtgagaaaagttttgtccatgccccagtcacatagtaaaagaaaactgaggtgcagaccggaggggttttcaaaggaa
 tcgattcttccaaaagggaatagtgcataagctcatcgctcgtaaaaaggactgggaccggaaaaagtagcgttggttcgtg
 agccctacagttgcctattctgtccttagtagtggtgcaaaagttgagaaggggaaaaatccaagaaactgaagtcagtc aaagaa
 ttattggggataacgattatggagcgctcgtcttttgaaaagaaccccatcgacttccttgaggcgaaagggttacaaggaa
 25 gtaaaaaaggatctcataattaaactaccaaagtatagtcgttttgagttagaaaaatggccgaaaaacggatggttggtagc
 gccagagagcttcaaaagggaacgaactcgactaccgtctaaatacgtgaatttctgtatttagcgtcccattacgag
 aagttgaaagggttcacctgaagataacgaacagaagcaactttttgttgagcagcacaacattatctcgacgaaatcata
 gagcaaatctcggaattcagtaagagagtcacctagctgatgccaatctggacaaagtattaagcgcatacaacaagcac
 agggataaaaccatacgtgagcagggcgaaaatattatccatttgtttactcttaccaacctcggcgctccagccgcattc
 30 aagtattttgacacaacgatagatcgcaaagagtacagatctaccaaggaggtgctagacgcgacactgattcaccaatcc
 atcacgggattatatgaaactcggatagatttgtcacagcttgggggtgac (SEQ ID NO: 36)

[0141] In embodiments, a nucleic acid encoding a Cas9 variant (SpCas9 D1135E) includes the nucleic acid sequence of SEQ ID NO: 37 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 37.

35 gataaaaagatttctattggttttagacatcggcactaattccgttggtgggtgtgcataaccgatgaatacaaagtacct
 tcaaagaaatttaagggtgttggggaacacagaccgtcattcgattaaaaagaatcttatcggtgcctcctattcgatagt

ggcgaaacggcagagggcactcgctgaaacgaaccgctcggagaaggtatacacgtcgcaagaaccgaatatgttactta
caagaaatTTTTtagcaatgagatggccaaagttgacgattctttctttcacggtttggaagagtccttccttgtcgaagag
gacaagaaacatgaacggcaccccatctttggaacatagtagatgaggtggcatatcatgaaaagtacccaacgatttat
cacctcagaaaaaagctagttgactcaactgataaagcggacctgaggttaatctacttggctcttgcccatatgataaag
5 ttccgtgggcactttctcattgaggggtgatctaaatccggacaactcggatgtcgacaaactgttcatccagtttagtaca
acctataatcagttgtttgaagagaaccctataaatgcaagtggcgtggatgcgaaggctattcttagcgcccgctctct
aaatcccgacggctagaaaacctgatcgacacaattaccggagagaagaaaaatgggttggttcggtaaccttatagcgctc
tactaggcctgacaccaaattttaagtcgaacttcgacttagctgaagatgccaaattgcagcttagtaaggacacgtac
gatgacgatctcgacaatctactggcacaattggagatcagtatgcggacttatttttggctgccaaaaaccttagcgat
10 gcaatcctcctatctgacatactgagagttaatactgagattaccaaggcgcggttatccgcttcaatgatcaaaagggtac
gatgaacatcaccaagacttgacacttctcaaggccctagtcggtcagcaactgcctgagaaatataaggaaatattcttt
gatcagtcgaaaaacgggtacgcaggttatattgacggcggagcaggtcaagaggaattctacaagtttatcaaaccata
ttagagaagatggatgggacggaagagttgcttgtaaaactcaatcggaagatctactgcgaagcagcggactttcgac
aacggtagcattccacatcaaatccacttaggcgaattgcatgctatacttagaaggcaggaggatttttatccgttcctc
15 aaagacaatcgtgaaaagattgagaaaatcctaaccctttcgcataccttactatgtgggacccctggcccgagggaactct
cggttcgcagtgatgacaagaaagtccgaagaaacgattactccatggaattttgaggaagttgtcgataaagggtgcgtca
gctcaatcgttcatcgagaggatgaccaactttgacaagaatttacccaacgaaaaagattatgcctaagcacagtttactt
tacgagtatttcacagtggtacaatgaactcacgaaagttaagtatgtcactgagggcatgcgtaaacccgcctttctaagc
ggagaacagaagaaagcaatagtagatctgttattcaagaccaaccgcaaagtgcaggttaagcaattgaaagaggactac
20 ttttaagaaaattgaatgcttcgattctgtcgagatctccggggtagaagatcgatttaatgcgtcacttggtacgtatcat
gacctcctaaagataattaaagataaggacttcctggataacgaagagaatgaagatatcttagaagatatagtgttgact
cttaccctctttgaagatcgggaaatgattgaggaaagactaaaaacatacgtcacctgttcgacgataagggttatgaaa
cagttaaagaggcgtcgctatacgggctggggacgattgtcgcggaacttatcaacgggataagagacaagcaaaagtgggt
aaaactattctcgattttctaaagagcgacggcttcgccaataggaactttatgcagctgatccatgatgactctttaacc
25 ttcaaagaggatatacaaaaaggcacaggtttccggacaaggggactcattgcacgaacatatgtcgaatcttgctggttcg
ccagccatcaaaaagggcatactccagacagtcaaagtagtggatgagctagttaagggtcatgggacgtcaciaaccggaa
aacattgtaatcgagatggcacgcgaaaatcaaacgactcagaaggggcaaaaaaacagtcgagagcggatgaagagaata
gaagaggggtattaaagaactgggcagccagatcttaaaggagcatcctgtggaaaatacccaattgcagaacgagaaactt
tacctctattacctacaaaatggaaggacatgtatgttgatcaggaactggacataaaacggtttatctgattacgacgtc
30 gatcacattgtacccaatcctttttgaaggacgattcaatcgacaataaagtgttacacgctcggataagaaccgaggg
aaaagtgcacaatgttccaagcgaggaagtcgtaagaaaaatgaagaactattggcggcagctcctaaatgcgaaactgata
acgcaaagaaagttcgataacttaactaaagctgagaggggtggcttgtctgaacttgacaaggccggattttattaaacgt
cagctcgtggaaacccgccaaatcaciaagcatgttgacagatactagattcccgaatgaatacgaaatacgacgagaaac
gataagctgattcgggaagtcaaagtaatcactttaaaagtcaaaattgggtgtcggacttcagaaaggattttcaattctat
35 aaagttaggagataaataactaccaccatgcgcacgacgcttatcttaatgccgtcgtagggaccgcactcattaagaaa
taccgaagctagaaagtgagtttgtgtatgggtgattacaaagtttatgacgtccgtaagatgatcgcgaaaagcgaaacag
gagataggcaaggctacagccaaatacttcttttattctaacattatgaatttctttaagacggaaatcactctggcaaac
ggagagatacgcaaacgacctttaattgaaaccaatggggagacaggtgaaatcgtatgggataaggccgggacttcgcg

acggtgagaaaagttttgtccatgccccagtcaacatagtaaagaaaactgaggtgcagaccggagggttttcaaaggaa
 tcgattcttccaaaaggaatagtataagctcatcgctcgtaaaaaggactgggacccgaaaaagtagcgtggcttcgag
 agccctacagttgcctattctgtcctagtagtggaaggttgagaagggaatccaagaaactgaagtcagtcaaagaa
 ttattggggataacgattatggagcgctcgtcttttgaaaagaaccccatcgacttccttgaggcgaaagggttacaaggaa
 5 gtaaaaaaggatctcataattaaactaccaaagtatagtcgtgtttgagttagaaaatggccgaaaacggatgttggttagc
 gccggagagcttcaaaagggaacgaactcgactaccgtctaaatacgtgaatttcctgtatttagcgtcccattacgag
 aagttgaaagggttcacctgaagataacgaacagaagcaactttttgttgagcagcacaacattatctcgacgaaatcata
 gagcaatttcggaattcagtaagagagtcctcctagctgatgccaatctggacaaagtattaagcgcatacaacaagcac
 aggataaacccatacgtgagcaggcggaatattatccatttgtttactcttaccacacctcggcgctccagccgcattc
 10 aagtattttgacacaacgatagatcgcaaacgatacacttctaccaaggaggtgctagacgcgacactgattcaccaatcc
 atcacgggattatatgaaactcggatagatttgtcacagcttgggggtgac (SEQ ID NO: 37)

[0142] In embodiments, a nucleic acid encoding a Cas9 variant (eSpCas9 1.1 mutant) includes the nucleic acid sequence of SEQ ID NO: 38 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 38.

15 GACAAGAAGTACAGCATCGGCCTGGACATCGGCACCAACTCTGTGGGCTGGGCCGTGATCACCGACGAGTACAAGGTGCCC
 AGCAAGAAATTCAGGTGCTGGGCAACACCGACCGGCACAGCATCAAGAAGAACCTGATCGGAGCCCTGCTGTTTCGACAGC
 GGCGAAACAGCCGAGGCCACCCGGCTGAAGAGAACCGCCAGAAGAAGATACACCAGACGGAAGAACCGGATCTGCTATCTG
 CAAGAGATCTTCAGCAACGAGATGGCCAAGGTGGACGACAGCTTCTTCACAGACTGGAAGAGTCCTTCCTGGTGGAAGAG
 GATAAGAAGCACGAGCGGCACCCCATCTTCGGCAACATCGTGGACGAGGTGGCCTACCACGAGAAGTACCCACCATCTAC
 20 CACCTGAGAAAGAACTGGTGGACAGCACCGACAAGGCCGACCTGCGGCTGATCTATCTGGCCCTGGCCACATGATCAAG
 TTCCGGGGCCACTTCCTGATCGAGGGCGACCTGAACCCCGACAACAGCGACGTGGACAAGCTGTTTCATCCAGCTGGTGCAG
 ACCTACAACCAGCTGTTTCGAGGAAAACCCCATCAACGCCAGCGGCGTGGACGCCAAGGCCATCCTGTCTGCCAGACTGAGC
 AAGAGCAGACGGCTGGAAAATCTGATCGCCCAGCTGCCC GGCGAGAAGAAGAATGGCCTGTTTCGGAACCTGATTGCCCTG
 AGCCTGGGCCTGACCCCAACTTCAAGAGCAACTTCGACCTGGCCGAGGATGCCAACTGCAGCTGAGCAAGGACACCTAC
 25 GACGACGACCTGGACAACCTGCTGGCCAGATCGGCGACCAAGTACGCCGACCTGTTTCTGGCCGCCAAGAACCTGTCCGAC
 GCCATCCTGCTGAGCGACATCCTGAGAGTGAACACCGAGATCACCAAGGCCCCCTGAGCGCCTCTATGATCAAGAGATAC
 GACGAGCACCACAGGACCTGACCCTGCTGAAAGCTCTCGTGCGGCAGCAGCTGCCTGAGAAGTACAAAGAGATTTTCTTC
 GACCAGAGCAAGAACGGCTACGCCGGCTACATTGACGGCGGAGCCAGCCAGGAAGAGTTCTACAAGTTCATCAAGCCCATC
 CTGGAAAAGATGGACGGCACCGAGGAAGTCTCGTGAAGCTGAACAGAGAGGACCTGCTGCGGAAGCAGCGGACCTTCGAC
 30 AACGGCAGCATCCCCACAGATCCACCTGGGAGAGCTGCACGCCATTCTGCGGCGGCAGGAAGATTTTTACCCATTCTG
 AAGGACAACCGGAAAAGATCGAGAAGATCCTGACCTTCGCATCCCCTACTACGTGGGCCCTCTGGCCAGGGGAAAACAGC
 AGATTGCGCTGGATGACCAGAAAGAGCGAGGAAACCATCACCCCTGGAACCTCGAGGAAGTGGTGGACAAGGGCGCTTCC
 GCCCAGAGCTTCATCGAGCGGATGACCAACTTCGATAAGAACCTGCCAACGAGAAGGTGCTGCCCAAGCACAGCCTGCTG
 TACGAGTACTTCACCGTGTATAACGAGCTGACCAAAGTGAATACGTGACCGAGGGAATGAGAAAGCCCGCCTTCCTGAGC
 35 GGCGAGCAGAAAAAGGCCATCGTGGACCTGCTGTTCAAGACCAACCGGAAAGTGACCGTGAAGCAGCTGAAAGAGGACTAC
 TTCAAGAAAATCGAGTGCTTCGACTCCGTGGAAATCTCCGGCGTGGAAGATCGGTTCAACGCCTCCCTGGGCACATACCAC

GATCTGCTGAAAATTATCAAGGACAAGGACTTCCTGGACAATGAGGAAAACGAGGACATTCTGGAAGATATCGTGCTGACC
 CTGACACTGTTTGAGGACAGAGAGATGATCGAGGAACGGCTGAAAACCTATGCCCACCTGTTTCGACGACAAAGTGATGAAG
 CAGCTGAAGCGGCGGAGATACACCGGCTGGGGCAGGCTGAGCCGGAAGCTGATCAACGGCATCCGGGACAAGCAGTCCGGC
 AAGACAATCCTGGATTTCTGAAGTCCGACGGCTTCGCCAACAGAACTTCATGCAGCTGATCCACGACGACAGCCTGACC
 5 TTTAAAGAGGACATCCAGAAAGCCCAGGTGTCCGGCCAGGGCGATAGCCTGCACGAGCACATTGCCAATCTGGCCGGCAGC
 CCCGCCATTAAAGAGGGCATCCTGCAGACAGTGAAGGTGGTGGACGAGCTCGTGAAAGTGATGGGCCGGCACAAGCCCGAG
 AACATCGTGATCGAAATGGCCAGAGAGAACCAGACCACCCAGAAGGGACAGAAGAAGAGCCGCGAGAGAATGAAGCGGATC
 GAAGAGGGCATCAAAGAGCTGGGCAGCCAGATCCTGAAAGAACACCCCGTGGAAAACACCCAGCTGCAGAACGAGAAGCTG
 TACCTGTACTACCTGCAGAATGGGCGGGATATGTACGTGGACCAGGAAGTGGACATCAACCGGCTGTCCGACTACGATGTG
 10 GACCATATCGTGCCTCAGAGCTTTCTGGCGGACGACTCCATCGACAACAAGGTGCTGACCAGAAGCGACAAGAACCGGGGC
 AAGAGCGACAACGTGCCCTCCGAAGAGGTCTGAAGAAGATGAAGAACTACTGGCGGCAGCTGCTGAACGCCAAGCTGATT
 ACCCAGAGAAAGTTCGACAATCTGACCAAGGCCGAGAGAGGCGGCCTGAGCGAACTGGATAAGGCCGGCTTCATCAAGAGA
 CAGCTGGTGGAAACCCGGCAGATCACAAGCACGTGGCACAGATCCTGGACTCCCGGATGAACACTAAGTACGACGAGAAT
 GACAAGCTGATCCGGGAAGTGAAAGTGATCACCTGAAGTCCAAGCTGGTGTCCGATTTCCGGAAGGATTTCCAGTTTTAC
 15 AAAGTGCGCGAGATCAACAACCTACCACCACGCCCACGACGCCTACCTGAACGCCGTCTGGGAACCGCCCTGATCAAAAAG
 TACCCTGCGCTGGAAAGCGAGTTCTGTACGGCGACTACAAGGTGTACGACGTGCGGAAGATGATCGCCAAGAGCGAGCAG
 GAAATCGGCAAGGCTACCGCCAAGTACTTCTTCTACAGCAACATCATGAACTTTTTCAAGACCGAGATTACCCTGGCCAAC
 GCGGAGATCCGGAAGGCGCCTCTGATCGAGACAAACGGCGAAACCGGGGAGATCGTGTGGGATAAGGGCCGGGATTTTGCC
 ACCGTGCGGAAAGTGCTGAGCATGCCCCAAGTGAATATCGTGAAAAAGACCGAGGTGCAGACAGGCGGCTTCAGCAAAGAG
 20 TCTATCCTGCCCAAGAGGAACAGCGATAAGCTGATCGCCAGAAAGAAGGACTGGGACCCTAAGAAGTACGGCGGCTTCGAC
 AGCCCCACCGTGGCCTATTCTGTGCTGGTGGTGGCCAAAGTGGAAGGGCAAGTCCAAGAACTGAAGAGTGTGAAAGAG
 CTGCTGGGGATCACCATCATGGAAGAAGCAGCTTCGAGAAGAATCCCATCGACTTTCTGGAAGCCAAGGGCTACAAAGAA
 GTGAAAAAGGACCTGATCATCAAGCTGCCTAAGTACTCCCTGTTTCGAGCTGGAAAACGGCCGGAAGAGAATGCTGGCCTCT
 GCCGGCGAACTGCAGAAGGGAAACGAACTGGCCCTGCCCTCCAAATATGTGAACTTCCTGTACCTGGCCAGCCACTATGAG
 25 AAGCTGAAGGGCTCCCCGAGGATAATGAGCAGAAACAGCTGTTTGTGGAACAGCACAAGCACTACCTGGACGAGATCATC
 GAGCAGATCAGCGAGTTCTCCAAGAGAGTGATCCTGGCCGACGCTAATCTGGACAAAGTGCTGTCCGCTACAACAAGCAC
 CGGGATAAGCCCATCAGAGAGCAGGCCGAGAATATCATCCACCTGTTTACCCTGACCAATCTGGGAGCCCCTGCCGCCTTC
 AAGTACTTTGACACCACCATCGACCGGAAGAGGTACACCAGCACCAAAGAGGTGCTGGACGCCACCCTGATCCACCAGAGC
 ATCACCGGCCTGTACGAGACACGGATCGACCTGTCTCAGCTGGGAGGCGACAAAAGGCCGGCGGCCACGAAAAAGGCCGGC
 30 CAGGCAAAAAAGAAAAAG (SEQ ID NO: 38)

[0143] In embodiments, a nucleic acid encoding a Cas9 variant (SpCas9 HF1) includes the nucleic acid sequence of SEQ ID NO: 39 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 39.

GATAAAAAGTATTCTATTGGTTTAGACATCGGCACTAATTCCGTTGGATGGGCTGTCATAACCGATGAATACAAAGTACCT
 35 TCAAAGAAATTTAAGGTGTTGGGGAACACAGACCGTCATTTCGATTAAAAAGAATCTTATCGGTGCCCTCCTATTTCGATAGT
 GGCGAAACGGCAGAGGCGACTCGCCTGAAACGAACCGCTCGGAGAAGGTATACACGTGCAAGAACCGAATATGTTACTTA

CAAGAAATTTTGTAGCAATGAGATGGCCAAAGTTGACGATTCTTTCTTTACCGTTTGGAAAGAGTCCTTCCTTGTGCAAGAG
GACAAGAAACATGAACGGCACCCCATCTTTGGAAACATAGTAGATGAGGTGGCATATCATGAAAAGTACCCAACGATTTAT
CACCTCAGAAAAAAGCTAGTTGACTCAACTGATAAAGCGGACCTGAGGTTAATCTACTTGGCTCTTGCCCATATGATAAAG
TTCCGTGGGCACCTTCTCATTGAGGGTGATCTAAATCCGGACAACCTCGGATGTGACAAAACCTGTTTCATCCAGTTAGTACAA
5 ACCTATAATCAGTTGTTTGAAGAGAACCCTATAAATGCAAGTGGCGTGGATGCGAAGGCTATTCTTAGCGCCCGCCTCTCT
AAATCCCGACGGCTAGAAAACCTGATCGCACAAATTACCCGGAGAGAAGAAAAATGGGTGTTGTCGGTAACCTTATAGCGCTC
TCACTAGGCCTGACACCAAATTTTAAGTCGAACTTCGACTTAGCTGAAGATGCCAAATTGCAGCTTAGTAAGGACACGTAC
GATGACGATCTCGACAATCTACTGGCACAATTTGGAGATCAGTATGCGGACTTATTTTTGGCTGCCAAAAACCTTAGCGAT
GCAATCCTCCTATCTGACATACTGAGAGTTAATACTGAGATTACCAAGGCGCCGTTATCCGCTTCAATGATCAAAAGGTAC
10 GATGAACATCACCAAGACTTGACACTTCTCAAGGCCCTAGTCCGTGAGCAACTGCCTGAGAAATATAAGGAAATATTCTTT
GATCAGTCGAAAAACGGGTACGCAGGTTATATTGACGGCGGAGCGAGTCAAGAGGAATTCTACAAGTTTATCAAACCCATA
TTAGAGAAGATGGATGGGACGGAAGAGTTGCTTGTA AAACTCAATCGCGAAGATCTACTGCGAAAGCAGCGGACTTTTCGAC
AACGGTAGCATTCCACATCAAATCCACTTAGGCGAATTGCATGCTATACTTAGAAGGCAGGAGGATTTTTATCCGTTCTCTC
AAAGACAATCGTGAAAAGATTGAGAAAATCCTAACCTTTTCGCATACCTTACTATGTGGGACCCCTGGCCCGAGGGAACTCT
15 CGGTTTCGCATGGATGACAAGAAAGTCCGAAGAAACGATTACTCCCTGGAATTTTGAGGAAGTTGTGCGATAAAGGTGCGTCA
GCTCAATCGTTCATCGAGAGGATGACCGCCTTTGACAAGAATTTACCGAACGAAAAAGTATTGCCTAAGCACAGTTTACTTT
TACGAGTATTTACAGTGTACAATGAACTCACGAAAGTTAAGTATGTCACTGAGGGCATGCGTAAACCCGCCTTTCTAAGC
GGAGAACAGAAGAAAGCAATAGTAGATCTGTTATTCAAGACCAACCGCAAAGTGACAGTTAAGCAATTGAAAGAGGACTAC
TTTAAGAAAATTGAATGCTTCGATTCTGTGAGATCTCCGGGGTAGAAGATCGATTTAATGCGTCACTTGGTACGTATCAT
20 GACCTCCTAAAGATAATTAAAGATAAGGACTTCCTGGATAACGAAGAGAATGAAGATATCTTAGAAGATATAGTGTTGACT
CTTACCCTCTTTGAAGATCGGGAAATGATTGAGGAAAGACTAAAAACATACGCTCACCTGTTTCGACGATAAGGTTATGAAA
CAGTTAAAGAGGCGTCGCTATACGGGCTGGGGAGCCTTGTCGCGGAACTTATCAACGGGATAAGAGACAAGCAAAGTGGT
AAAATATTCTCGATTTTCTAAAGAGCGACGGCTTCGCCAATAGGAACTTTATGGCCCTGATCCATGATGACTCTTTAACC
TTCAAAGAGGATATACAAAAGGCACAGGTTTCCGGACAAGGGGACTCATTGCACGAACATATTGCGAATCTTGCTGGTTTCG
25 CCAGCCATCAAAAAGGGCATACTCCAGACAGTCAAAGTAGTGGATGAGCTAGTTAAGGTCATGGGACGTACAAAACCGGAA
AACATTGTAATCGAGATGGCACGCGAAAATCAAACGACTCAGAAGGGGCAAAAAACAGTCGAGAGCGGATGAAGAGAATA
GAAGAGGGTATTAAAGAACTGGGCAGCCAGATCTTAAAGGAGCATCCTGTGGAATAACCCAATTGCAGAACGAGAACTTT
TACCTCTATTACCTACAAAATGGAAGGGACATGTATGTTGATCAGGAACTGGACATAAACCGTTTATCTGATTACGACGTC
GATCACATTGTACCCCAATCCTTTTTGAAGGACGATTCAATCGACAATAAAGTGCTTACACGCTCGGATAAGAACCGAGGG
30 AAAAGTGACAATGTTCCAAGCGAGGAAGTCGTAAAGAAAATGAAGAACTATTGGCGGCAGCTCCTAAATGCGAACTGATA
ACGCAAAGAAAGTTTCGATAACTTAACTAAAGCTGAGAGGGGTGGCTTGTCTGAACTTGACAAGGCCGGATTTATTAAACGT
CAGCTCGTGGAAACCCGCGCCATCACAAGCATGTTGCCAGATACTAGATTCCCGAATGAATACGAAATACGACGAGAAC
GATAAGCTGATTTCGGGAAGTCAAAGTAATCACTTTAAAGTCAAATTTGGTGTGCGACTTCAGAAAGGATTTTCAATTCTAT
AAAGTTAGGGAGATAAATACTACCACCATGCGCACGACGCTTATCTTAATGCCGTGCTAGGGACCGCACTCATTAAGAAA
35 TACCCGAAGCTAGAAAGTGAGTTTGTGTATGGTGATTACAAAGTTTATGACGTCCGTAAGATGATCGCGAAAAGCGAACAG
GAGATAGGCAAGGCTACAGCCAAATACTTCTTTTATTCTAACATTATGAATTTCTTTAAGACGGAAATCACTCTGGCAAAC
GGAGAGATACGCAAACGACCTTTAATTGAAACCAATGGGGAGACAGGTGAAATCGTATGGGATAAGGGCCGGGACTTCGCG
ACGGTGAGAAAAGTTTTGTCCATGCCCCAAGTCAACATAGTAAGAAAACCTGAGGTGCAGACCGGAGGGTTTTCAAAGGAA

TCGATTCTTCCAAAAGGAATAGTGATAAGCTCATCGCTCGTAAAAAGGACTGGGACCCGAAAAAGTACGGTGGCTTCGAT
 AGCCCTACAGTTGCCTATTCTGTCTAGTAGTGGCAAAAGTTGAGAAGGGAAAAATCCAAGAACTGAAGTCAGTCAAAGAA
 TTATTGGGGATAACGATTATGGAGCGCTCGTCTTTTGAAGAAGACCCCATCGACTTCCTTGAGGCGAAAGGTTACAAGGAA
 GTAAAAAAGGATCTCATAATTAACTACCAAAGTATAGTCTGTTTGAGTTAGAAAATGGCCGAAAACGGATGTTGGCTAGC
 5 GCCGGAGAGCTTCAAAGGGGAACGAACTCGCACTACCGTCTAAATACGTGAATTTCTGTATTTAGCGTCCCATTACGAG
 AAGTTGAAAGGTTACCTGAAGATAACGAACAGAAGCAACTTTTTGTTGAGCAGCACAAACATTATCTCGACGAAATCATA
 GAGCAAATTTCCGAATTCAGTAAGAGAGTCATCCTAGCTGATGCCAATCTGGACAAAGTATTAAGCGCATACAACAAGCAC
 AGGGATAAACCCATACGTGAGCAGGCGGAAAATATTATCCATTTGTTTACTCTTACCAACCTCGGCGCTCCAGCCGCATTC
 AAGTATTTTGACACAACGATAGATCGCAAACGATACACTTCTACCAAGGAGGTGCTAGACGCGACACTGATTACCAATCC
 10 ATCAGGGGATTATATGAAACTCGGATAGATTTGTACAGCTTGGGGGTGAC (SEQ ID NO: 39)

[0144] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to Cas9, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes SEQ ID NO: 2, and the nucleic acid encoding the Cas9 protein includes SEQ ID NO: 32.

[0145] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to Cas9, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2, and the nucleic acid encoding the Cas9 protein includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 32.

[0146] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to Cas9, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes SEQ ID NO: 2, and the nucleic acid encoding the Cas9 protein includes SEQ ID NO: 33.

[0147] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to Cas9, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2, and the nucleic acid encoding the Cas9 protein includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 33.

[0148] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to Cpf1, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes SEQ ID NO: 2, and the nucleic acid encoding the Cpf1 protein includes SEQ ID NO: 34.

[0149] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to Cpf1, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2, and the nucleic acid encoding the Cpf1 protein includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 34.

[0150] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes SEQ ID NO: 35.

[0151] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial
5 localization amino acid sequence includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 35.

[0152] In embodiments, the composition described herein includes a delivery vehicle (e.g., a
10 nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes SEQ ID NO: 36.

[0153] In embodiments, the composition described herein includes a delivery vehicle (e.g., a
15 nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes a nucleic acid sequence having at least 80%, 85%, 90%,
20 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 36.

[0154] In embodiments, the composition described herein includes a delivery vehicle (e.g., a
25 nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes SEQ ID NO: 37.

[0155] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial
5 localization amino acid sequence includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 37.

[0156] In embodiments, the composition described herein includes a delivery vehicle (e.g., a
10 nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes SEQ ID NO: 38.

[0157] In embodiments, the composition described herein includes a delivery vehicle (e.g., a
15 nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes a nucleic acid sequence having at least 80%, 85%, 90%,
20 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 38.

[0158] In embodiments, the composition described herein includes a delivery vehicle (e.g., a
25 nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial localization amino acid sequence includes SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes SEQ ID NO: 39.

[0159] In embodiments, the composition described herein includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant, where the nucleic acid is bound to the delivery vehicle, the nucleic acid encoding the mitochondrial
5 localization amino acid sequence includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2, and the nucleic acid encoding the Cas9 variant protein includes a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 39.

[0160] In another aspect, there is provided a protein including a mitochondrial localization amino
10 acid sequence covalently attached to Cpf1 or a Class II CRISPR endonuclease or a Cas9 variant.

[0161] In embodiments, the protein includes a mitochondrial localization amino acid sequence covalently attached to Cpf1.

[0162] In embodiments, the protein includes a mitochondrial localization amino acid sequence covalently attached to a Class II CRISPR endonuclease.

15 [0163] In embodiments, the protein includes a mitochondrial localization amino acid sequence covalently attached to a Cas9 variant.

[0164] In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including SEQ ID NO: 1 and Cpf1 protein including SEQ ID NO: 25.

[0165] In embodiments, the protein described herein includes a mitochondrial localization amino
20 acid sequence including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and Cpf1 protein including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 25.

[0166] In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including SEQ ID NO: 1 and a Cas9 variant protein including SEQ ID NO: 26.

25 [0167] In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and a Cas9 variant protein including an amino acid

sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 26.

[0168] In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including SEQ ID NO: 1 and a Cas9 variant protein including SEQ ID NO: 27.

5 **[0169]** In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and a Cas9 variant protein including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 27.

10 **[0170]** In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including SEQ ID NO: 1 and a Cas9 variant protein including SEQ ID NO: 28.

[0171] In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and a Cas9 variant protein including an amino acid
15 sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 28.

[0172] In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including SEQ ID NO: 1 and a Cas9 variant protein including SEQ ID NO: 29.

[0173] In embodiments, the protein described herein includes a mitochondrial localization amino
20 acid sequence including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and a Cas9 variant protein including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 29.

[0174] In embodiments, the protein described herein includes a mitochondrial localization amino
25 acid sequence including SEQ ID NO: 1 and a Cas9 variant protein including SEQ ID NO: 30.

[0175] In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and a Cas9 variant protein including an amino acid

sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 30.

[0176] In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including SEQ ID NO: 1 and a Cas9 variant protein including SEQ ID NO: 31.

5 **[0177]** In embodiments, the protein described herein includes a mitochondrial localization amino acid sequence including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 1 and a Cas9 variant protein including an amino acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 31.

10 **[0178]** In another aspect, there is provided a nucleic acid encoding a protein as disclosed hereinabove. In embodiments, the nucleic acid encodes a protein including a mitochondrial localization amino acid sequence covalently attached to Cpf1 or a Class II CRISPR endonuclease or a Cas9 variant.

[0179] In embodiments, the nucleic acid described herein encodes a protein including a
15 mitochondrial localization amino acid sequence covalently attached to Cpf1.

[0180] In embodiments, the nucleic acid described herein encodes a protein including a mitochondrial localization amino acid sequence covalently attached to a Class II CRISPR endonuclease.

[0181] In embodiments, the nucleic acid described herein encodes a protein including a
20 mitochondrial localization amino acid sequence covalently attached to a Cas9 variant.

[0182] In embodiments, the nucleic acid described herein includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including SEQ ID NO: 32.

[0183] In embodiments, the nucleic acid described herein includes a nucleic acid including a
25 nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 32.

[0184] In embodiments, the nucleic acid described herein includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including SEQ ID NO: 33.

5 [0185] In embodiments, the nucleic acid described herein includes a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 33.

[0186] In embodiments, the nucleic acid described herein includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including SEQ ID NO: 34.

10 [0187] In embodiments, the nucleic acid described herein includes a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 34.

[0188] In embodiments, the nucleic acid described herein includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including SEQ ID NO: 35.

15 [0189] In embodiments, the nucleic acid described herein includes a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 35.

20 [0190] In embodiments, the nucleic acid described herein includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including SEQ ID NO: 36.

[0191] In embodiments, the nucleic acid described herein includes a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 36.

25 [0192] In embodiments, the nucleic acid described herein includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including SEQ ID NO: 37.

[0193] In embodiments, the nucleic acid described herein includes a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 37.

5 **[0194]** In embodiments, the nucleic acid described herein includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including SEQ ID NO: 38.

[0195] In embodiments, the nucleic acid described herein includes a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 38.

[0196] In embodiments, the nucleic acid described herein includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including SEQ ID NO: 39.

15 **[0197]** In embodiments, the nucleic acid described herein includes a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 39.

[0198] In another aspect, there is provided a nucleic acid including a mitochondrial import sequence and a single-guide RNA (sgRNA) sequence.

20 **[0199]** In embodiments, the mitochondrial import sequence includes a D loop, an F loop, an MRP loop, an RNP loop, a γ 5s loop, or any combination thereof.

[0200] In embodiments, the mitochondrial import sequence includes SEQ ID NO: 3 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 3. In embodiments, the mitochondrial import sequence includes SEQ ID NO: 4 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 4. In
25 embodiments, the mitochondrial import sequence includes SEQ ID NO: 5 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 5. In embodiments, the mitochondrial import sequence includes SEQ ID NO: 6 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 6. In

embodiments, the mitochondrial import sequence includes SEQ ID NO: 7 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 7. In

embodiments, the mitochondrial import sequence includes any combination of SEQ ID NOs: 3-7 or any combination of the nucleic acid sequences having at least 80%, 85%, 90%, 95%, 96%, 97%,

5 98%, 99% or 100% identity to any one of SEQ ID Nos 3-7.

[0201] In embodiments, the sgRNA sequence includes a guide sequence (i.e., a nucleic acid sequence that is complementary to a mitochondrial DNA (mtDNA) target sequence). In embodiments, the mtDNA target sequence includes at least one mutation or deletion. In embodiments, the mtDNA target sequence is 8-100 nucleotides in length.

10 **[0202]** In embodiments, the mtDNA target sequence may optionally have a minimum length of one of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides.

15 **[0203]** In embodiments, the mtDNA target sequence may optionally have a maximum length of one of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides.

20 **[0204]** In embodiments, the mtDNA target sequence may optionally have a length of 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, or 100 nucleotides.

25 **[0205]** Where the mtDNA target sequence is less than 100 (99 or less) nucleotides in length, the sequence is one of 99, 98, 97, 96, 95, 94, 93, 92, 91, 90, 89, 88, 87, 86, 85, 84, 83, 82, 81, 80, 79, 78, 77, 76, 75, 74, 73, 72, 71, 70, 69, 68, 67, 66, 65, 64, 63, 62, 61, 60, 59, 58, 57, 56, 55, 54, 53, 52, 51, 50, 49, 48, 47, 46, 45, 44, 43, 42, 41, 40, 39, 38, 37, 36, 35, 34, 33, 32, 31, 30, 29, 28, 27, 26, 25, 24, 23, 22, 21, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9 or 8 nucleotides in length. In embodiments,

the mtDNA target sequence is less than 90 nucleotides in length. In embodiments, the mtDNA target sequence is less than 80 nucleotides in length. In embodiments, the mtDNA target sequence is less than 70 nucleotides in length. In embodiments, the mtDNA target sequence is less than 60 nucleotides in length. In embodiments, the mtDNA target sequence is less than 50 nucleotides in length. In embodiments, the mtDNA target sequence is less than 40 nucleotides in length. In embodiments, the mtDNA target sequence is less than 30 nucleotides in length. In embodiments, the mtDNA target sequence is less than 20 nucleotides in length. In embodiments, the mtDNA target sequence is less than 10 nucleotides in length.

[0206] In embodiments, the mtDNA target sequence is between 10 and 99 nucleotides in length.

In embodiments, the mtDNA target sequence is between 10 and 80 nucleotides in length. In embodiments, the mtDNA target sequence is between 10 and 70 nucleotides in length. In embodiments, the mtDNA target sequence is between 10 and 60 nucleotides in length. In embodiments, the mtDNA target sequence is between 10 and 50 nucleotides in length. In embodiments, the mtDNA target sequence is between 10 and 40 nucleotides in length. In embodiments, the mtDNA target sequence is between 10 and 30 nucleotides in length. In embodiments, the mtDNA target sequence is between 10 and 25 nucleotides in length. In embodiments, mtDNA target sequence is between 16 and 24 nucleotides in length.

[0207] Exemplary mt DNA target sequence may includes, but is not limited to, a fragement of any one of SEQ ID Nos: 8-22.

[0208] In embodiments, the sgRNA sequence comprises a transactivating crRNA (tracrRNA) sequence. In embodiments, the sgRNA sequence includes a crRNA sequence (i.e., a guide sequence and a tracr mate sequence). In embodiments, the sgRNA sequence includes a tracrRNA sequence and a crRNA sequence.

[0209] Exemplary nucleic acid sequences including a mitochondrial import sequence and a

sgRNA sequence include, but are not limited to, SEQ ID Nos: 40-48. In embodiments, an extra “G” may be added in front of sgRNAs that do not start with a G nucleotide in order to obtain efficient expression from U6 promoter.

Table 5		
Mito-sgRNA	Sequence	SEQ ID NO
sgRNA-D	GTTT TAGAGCTAGAAATAGCAAGTTAAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCggga GCGCAATCGG TAGCGC ttcccTTTTT	40
sgRNA-F	GTTT TAGAGCTAGAAATAGCAAGTTAAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCggga GAGCCCCCTACAG GGCTC ttcccTTTTT	41
sgRNA-DF	GTTT TAGAGCTAGAAATAGCAAGTTAAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCggga GCGCAATCGGTAG CGCGAGCCCCCTACAGGGCTC ttcccTTTTT	42
sgRNA-zDF	GTTT TAGAGCTAG GGCCGCGCAATCGGTAG CGCGGCCTAGCAAGTTAAAATAAGGCTAG TCCGTTATCAACTT TGGCCGAGCCCCCTAC AGGGCTCGGCCAAGTGGCACCGAGTCGGT GCTTTTT	43
sgRNA-RNP (RNase P)	GTTT TAGAGCTAGAAATAGCAAGTTAAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCggga TCTCCCTGAG CTTCAGGGAG ttcccTTTTT	44
sgRNA-MRP	GTTT TAGAGCTAGAAATAGCAAGTTAAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCggga AGAAGCGTAT CCCGCTGAGC ttcccTTTTT	45
sgRNA-g5s (gamma 5s)	GTTT TAGAGCTAGAAATAGCAAGTTAAAAT AAGGCTAGTCCGTTATCAACTTGAAAAAGT GGCACCGAGTCGGTGCggga GGCCTGGTTA GTA CTTGGATGGGAGACCGCCAAGGAAT ACCGGGTG ttcccTTTTT	46
Cpf1 sgRNA	TAATTTCTACTCTTGTAGATNNNNNNNNNN NNNNNNNNNNNN	47
saCas9 sgRNA	NNNNNNNNNNNNNNNNNNNNNNNNNGTTT TACTCTGGAAACAGAATCTACTAAAACAA GGCAAAATGCCGTGTTTATCTCGTCAACTT GTTGGCGAGA	48
Capital: crRNA and tracrRNA Capital and bold: mitochondrial import sequence Non-capital: stem for structural stability of mitoloops		

[0210] In embodiments, the sgRNA sequence may include a nuclear-encoded and mitochondrial-localizing tRNA sequence. A transfer RNA (abbreviated tRNA and archaically referred to as sRNA, for soluble RNA) is an adaptor molecule composed of RNA, typically 76 to 90 nucleotides in length, that serves as the physical link between the mRNA and the amino acid sequence of proteins. It does this by carrying an amino acid to the protein synthetic machinery of a cell (ribosome) as directed by a three-nucleotide sequence (codon) in a messenger RNA (mRNA). As such, tRNAs are a necessary component of translation, the biological synthesis of new proteins according to the genetic code. a nuclear-encoded and mitochondrial-localizing tRNA sequence refers to a tRNA has tRNA genes in its nuclear genome, but it is later transported into mitochondria to function after it is transcribed and matured.

[0211] Exemplary nuclear-encoded and mitochondrial-localizing tRNA sequences may include, but are not limited to, any one of SEQ ID Nos: 49-53.

Table 6		
5' addition of tRNA to Cpf1 sgRNA	Sequence	SEQ ID NO
Rat tRNA CUG1 Gln	GGTCCCATGGTGTAAATGGTTAGCACTCTGG ACTCTGAATCCAGCGATCCGAGTTCAAATC TCGGTGGGACCTCCA	49
Rat tRNA CUG2 Gln	GGTCCCATGGTGTAAATGGTGAGCACTCTGG ACTCTGAATCCAGCGATCCGAGTTCAAATC TCGGTGGGACCTCCA	50
Rat tRNA UUG1 Gln	GGTCCCATGGTGTAAATGGTTAGCACTCTGG ACTTTGAATCCAGCGATCCGAGTTCAAATC TCGGTGGGACCTCCA	51
Rat tRNA UUG2 Gln	GGTCCCATGGTGTAAATGGTGAGCACTCTGG ACTTTGAATCCAGCGATCCGAGTTCAAATC TCGGTGGGACCTCCA	52
human tRNA CUG Gln	ggtccatggtgtaatggtgagcactctggactctgaatccagcgatcc gagttcgagtctcggtggaacctCCA	53

[0212] In embodiments, the nucleic acid described herein is bound to a delivery vehicle. In embodiments, the delivery vehicle is a nanoparticle, a lipid particle or a viral vector.

[0213] In embodiments, the nucleic acid is DNA. In embodiments, the nucleic acid is RNA. In embodiments, the RNA includes one or more modified nucleotides or nucleotide analogues.

[0214] In another aspect, there is provided a complex that includes a protein including a mitochondrial localization amino acid sequence described herein (e.g., SEQ ID NO: 1) covalently attached to an RNA-guided DNA endonuclease enzyme described herein (e.g., Cas9, Cpf1, a Class II CRISPR endonuclease or a Cas9 variant) and a nucleic acid including one or more mitochondrial import sequences described herein (e.g., SEQ ID Nos: 3-7) and a sgRNA sequence described herein, where the protein binds to the nucleic acid.

[0215] In another aspect, there is provided a vector including any nucleic acid disclosed herein. In embodiments, the vector includes a nucleic acid according to any embodiments set forth above, or a combination thereof.

[0216] In embodiments, the vector includes a nucleic acid encoding a protein that includes a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme (e.g., Cas9, Cpf1, a Class II CRISPR endonuclease or a Cas9 variant).

[0217] In embodiments, the vector includes a nucleic acid including SEQ ID NO: 2 and a nucleic acid including any one of SEQ ID NOs: 32-39.

[0218] In embodiments, the vector includes a nucleic acid including a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 2 and a nucleic acid including a nucleic acid sequence having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to any one of SEQ ID NOs: 32-39.

[0219] In embodiments, the vector includes a nucleic acid including one or more (e.g., 1, 2, 3, 4 or more) mitochondrial import sequences and a sgRNA sequence.

[0220] In embodiments, the vector includes a nucleic acid including at least one (e.g., 1, 2, 3, 4 or more) mitochondrial import sequence of any one of SEQ ID Nos: 3-7 and a sgRNA sequence.

[0221] In embodiments, the vector is a replication incompetent viral vector. In embodiments, the replication incompetent viral vector is a replication incompetent lentiviral, adeno-associated viral, or adenoviral vector.

[0222] In another aspect, there is provided a pharmaceutical composition including a composition disclosed herein, a protein disclosed herein, a nucleic acid disclosed herein, a vector disclosed herein, and a pharmaceutically acceptable excipient.

[0223] "Pharmaceutically acceptable excipient" and "pharmaceutically acceptable carrier" refer to a substance that aids the administration of an active agent to and absorption by a subject and can be included in the compositions of the present invention without causing a significant adverse toxicological effect on the patient. Non-limiting examples of pharmaceutically acceptable excipients include water, NaCl, normal saline solutions, lactated Ringer's, normal sucrose, normal glucose, binders, fillers, disintegrants, lubricants, coatings, sweeteners, flavors, salt solutions (such as Ringer's solution), alcohols, oils, gelatins, carbohydrates such as lactose, amylase or starch, fatty acid esters, hydroxymethylcellulose, polyvinyl pyrrolidone, and colors, and the like. Such preparations can be sterilized and, if desired, mixed with auxiliary agents such as lubricants, preservatives, stabilizers, wetting agents, emulsifiers, salts for influencing osmotic pressure, buffers, coloring, and/or aromatic substances and the like that do not deleteriously react with the compounds of the invention. One of skill in the art will recognize that other pharmaceutical excipients are useful in the present invention.

[0224] The term "pharmaceutically acceptable salt" refers to salts derived from a variety of organic and inorganic counter ions well known in the art and include, by way of example only, sodium, potassium, calcium, magnesium, ammonium, tetraalkylammonium, and the like; and when the molecule contains a basic functionality, salts of organic or inorganic acids, such as hydrochloride, hydrobromide, tartrate, mesylate, acetate, maleate, oxalate and the like.

[0225] The term "preparation" is intended to include the formulation of the active compound with encapsulating material as a carrier providing a capsule in which the active component with or without other carriers, is surrounded by a carrier, which is thus in association with it. Similarly, cachets and lozenges are included. Tablets, powders, capsules, pills, cachets, and lozenges can be used as solid dosage forms suitable for oral administration.

[0226] The pharmaceutical preparation is optionally in unit dosage form. In such form the preparation is subdivided into unit doses containing appropriate quantities of the active component. The unit dosage form can be a packaged preparation, the package containing discrete quantities of preparation, such as packeted tablets, capsules, and powders in vials or ampoules. Also, the unit dosage form can be a capsule, tablet, cachet, or lozenge itself, or it can be the appropriate number of any of these in packaged form. The unit dosage form can be of a frozen dispersion.

III. Methods

[0227] In another aspect, there is provided a method of altering expression of at least one mitochondrial nucleic acid sequence, the method including introducing into an eukaryotic cell a nucleic acid as disclosed herein, where the nucleic acid includes a mitochondrial import sequence
5 and a single-guide RNA (sgRNA) sequence.

[0228] In embodiments, the mitochondrial import sequence includes a D loop, an F loop, an MRP loop, an RNP loop, a γ 5s loop, or any combination thereof. In embodiments, the mitochondrial import sequence includes SEQ ID NO: 3 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 3. In embodiments, the mitochondrial
10 import sequence includes SEQ ID NO: 4 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 4. In embodiments, the mitochondrial import sequence includes SEQ ID NO: 5 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 5. In embodiments, the mitochondrial import sequence includes SEQ ID NO: 6 or a nucleic acid having at least 80%, 85%, 90%, 95%,
15 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 6. In embodiments, the mitochondrial import sequence includes SEQ ID NO: 7 or a nucleic acid having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to SEQ ID NO: 7. In embodiments, the mitochondrial import sequence includes any combination of SEQ ID NOs: 3-7 or any combination of the nucleic acid sequences having at least 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% or 100% identity to any
20 one of SEQ ID Nos 3-7.

[0229] In embodiments, the sgRNA sequence includes a guide sequence (i.e., a nucleic acid sequence that is complementary to a mitochondrial DNA (mtDNA) target sequence). In embodiments, the mtDNA target sequence includes at least one mutation or deletion. In embodiments, the mtDNA target sequence is 8-100 nucleotides in length.

25 [0230] Exemplary mt DNA target sequence may include, but is not limited to, a fragment (e.g., 10-30 nucleotides) of any one of SEQ ID Nos: 8-22.

[0231] Exemplary guide sequence may include, but is not limited to, any one of SEQ ID Nos: 55-105.

[0232] In embodiments, the sgRNA sequence comprises a transactivating crRNA (tracrRNA) sequence. In embodiments, the sgRNA sequence includes a crRNA sequence (i.e., a guide sequence and a tracr mate sequence). In embodiments, the sgRNA sequence includes a tracrRNA sequence and a crRNA sequence. Exemplary crRNA sequence with or without a tracrRNA sequence includes, but is not limited to, any one of SEQ ID NOs: 106-108.

[0233] Exemplary nucleic acid sequences including a mitochondrial import sequence and a sgRNA sequence include, but are not limited to, SEQ ID Nos: 40-48.

[0234] In embodiments, the sgRNA sequence may include a nuclear-encoded and mitochondrial-localizing tRNA sequence. Exemplary nuclear-encoded and mitochondrial-localizing tRNA sequences may include, but are not limited to, any one of SEQ ID Nos: 49-53.

[0235] In embodiments, the method further includes introducing into the eukaryotic cell a composition as disclosed herein, a protein as disclosed herein, or another nucleic acid disclosed herein. For example, the composition includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein including a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme (e.g., Cas9, Cpf1, a Class II CRISPR endonuclease or a Cas9 variant), where the protein is bound to the delivery vehicle. For example, the protein includes a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme (e.g., Cas9, Cpf1, a Class II CRISPR endonuclease or a Cas9 variant). For example, another nucleic acid includes a nucleic acid encoding a protein including a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme (e.g., Cas9, Cpf1, a Class II CRISPR endonuclease or a Cas9 variant).

[0236] In embodiments, the eukaryotic cell is an oocyte. In embodiments, the eukaryotic cell is part of a fertilized embryo.

[0237] In embodiments, the mitochondrial nucleic acid sequence comprises at least one mutation or deletion. Exemplary mitochondrial nucleic acid sequences including at least one mutation or deletion include, but are not limited to, any one of SEQ ID Nos: 8-22.

[0238] Altering expression refers to an increase of expression of a wild type mitochondrial nucleic acid sequence, a decrease of expression of a mutant mitochondrial nucleic acid sequence or a combination thereof. In embodiments, altering expression of at least one mitochondrial nucleic acid

sequence refers to a decrease (i.e., downregulation) of at least one mutant mitochondrial nucleic acid sequence.

[0239] Levels of expression may be quantitated for absolute comparison, or relative comparisons may be made.

5 [0240] In some embodiments upregulation of expression may be considered to be present when the level of expression in the test sample is at least 1.1 times that of a reference level. More preferably, the level of expression may be selected from one of at least 1.2, at least 1.3, at least 1.4, at least 1.5, at least 1.6, at least 1.7, at least 1.8, at least 1.9, at least 2.0, at least 2.1, at least 2.2, at least 2.3, at least 2.4 at least 2.5, at least 2.6, at least 2.7, at least 2.8, at least 2.9, at least 3.0, at least 3.5, at least
10 4.0, at least 5.0, at least 6.0, at least 7.0, at least 8.0, at least 9.0, or at least 10.0 times that of the reference level.

[0241] In some embodiments downregulation of expression may be considered to be present when the level of expression in the test sample is at least 1.1 times lower than that of a reference level. More preferably, the level of expression may be selected from one of at least 1.2, at least 1.3, at least
15 1.4, at least 1.5, at least 1.6, at least 1.7, at least 1.8, at least 1.9, at least 2.0, at least 2.1, at least 2.2, at least 2.3, at least 2.4 at least 2.5, at least 2.6, at least 2.7, at least 2.8, at least 2.9, at least 3.0, at least 3.5, at least 4.0, at least 5.0, at least 6.0, at least 7.0, at least 8.0, at least 9.0, or at least 10.0 times lower than that of the reference level.

[0242] Expression levels may be determined by one of a number of known *in vitro* assay
20 techniques, such as PCR based assays, in situ hybridization assays, flow cytometry assays, immunological or immunohistochemical assays.

[0243] In embodiments, the method may further comprise a step of determining expression level (pre-treated level) of at least one mitochondrial nucleic acid sequence before introducing the nucleic acid into the eukaryotic cell. In embodiments, the method may further comprise another step of
25 determining expression level (post-treated level) of the same at least one mitochondrial nucleic acid sequence after introducing the nucleic acid into the eukaryotic cell. In embodiments, the method may further comprise a step of comparing the difference of the pre-treated level and the post-treated level in order to determine if the expression of the at least one mitochondrial nucleic acid sequence of interest has been altered.

[0244] In another aspect, there is provided a method of treating a mitochondrial disorder in a subject in need thereof, the method including administering to the subject an effective amount of a composition disclosed herein and an effective amount of a nucleic acid disclosed herein. For example, the composition includes a delivery vehicle (e.g., a nanoparticle, a lipid particle or a viral vector) and a protein including a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme (e.g., Cas9, Cpf1, a Class II CRISPR endonuclease or a Cas9 variant), where the protein is bound to the delivery vehicle. For example, the nucleic acid includes a mitochondrial import sequence and a single-guide RNA (sgRNA) sequence.

[0245] In another aspect, there is provided a method of treating a mitochondrial disorder in a subject in need thereof, the method including administering to the subject an effective amount of a protein disclosed herein and an effective amount of a nucleic acid disclosed herein. For example, the protein includes a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme (e.g., Cas9, Cpf1, a Class II CRISPR endonuclease or a Cas9 variant). For example, the nucleic acid includes a mitochondrial import sequence and a single-guide RNA (sgRNA) sequence.

[0246] In another aspect, there is provided a method of treating a mitochondrial disorder in a subject in need thereof, the method including administering to the subject an effective amount of a first nucleic acid disclosed herein and an effective amount of second nucleic acid disclosed herein. For example, the first nucleic acid includes a nucleic acid encoding a protein including a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme (e.g., Cas9, Cpf1, a Class II CRISPR endonuclease or a Cas9 variant). For example, the second nucleic acid includes a mitochondrial import sequence and a single-guide RNA (sgRNA) sequence.

[0247] Further to aspect or embodiments disclosed herein providing a method of treating a mitochondrial disorder in a subject in need thereof, in embodiments the mitochondrial disorder is selected from the group consisting of Myoclonic Epilepsy with Ragged Red Fibers (MERRF); Mitochondrial Myopathy, Encephalopathy, Lactacidosis, and Stroke (MELAS); Maternally Inherited Diabetes and Deafness (MIDD); Leber's Hereditary Optic Neuropathy (LHON); chronic progressive external ophthalmoplegia (CPEO); Leigh Disease; Kearns-Sayre Syndrome (KSS); Friedreich's Ataxia (FRDA); Co-Enzyme Q10 (CoQ10) Deficiency; Complex I Deficiency; Complex II

Deficiency; Complex III Deficiency; Complex IV Deficiency; Complex V Deficiency; other myopathies; cardiomyopathy; encephalomyopathy; renal tubular acidosis; neurodegenerative diseases; Parkinson's disease; Alzheimer's disease; amyotrophic lateral sclerosis (ALS); motor neuron diseases; hearing and balance impairments; or other neurological disorders; epilepsy; genetic diseases; Huntington's Disease; mood disorders; nucleoside reverse transcriptase inhibitors (NRTI) treatment; HIV-associated neuropathy; schizophrenia; bipolar disorder; age-associated diseases; cerebral vascular diseases; macular degeneration; diabetes; and cancer.

IV. Kits

[0248] In another aspect, there is provided a kit including a first nucleic acid disclosed herein, and a protein described herein that includes a mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA endonuclease enzyme or a second nucleic acid sequence encoding the protein. For example, the first nucleic acid includes a mitochondrial import sequence and a single-guide RNA (sgRNA) sequence.

[0249] In embodiments, the mitochondrial localization amino acid sequence is N-terminal to the RNA-guided DNA endonuclease enzyme. In embodiments, the mitochondrial localization amino acid sequence is a cytochrome c oxidase subunit VIII (Cox8) sequence. In embodiments, the RNA-guided DNA endonuclease enzyme is Cas9 or Cpf1 or a Class II CRISPR endonuclease. In embodiments, the RNA-guided DNA endonuclease enzyme has no nuclear localization sequence. In embodiments, the Cas9 is a mutant Cas9, wherein the mutant Cas9 has one or more mutations that increase its binding specificity to PAM compared to wild type Cas9.

[0250] In embodiments, the first nucleic acid sequence forms part of a first viral vector nucleic acid sequence. In embodiments, the second nucleic acid sequence forms part of a second viral vector nucleic acid sequence. In embodiments, the first viral vector and the second viral vector are the same. In embodiments, the first viral vector is different from the second viral vector.

V. Examples

[0251] The following examples illustrate certain specific embodiments of the invention and are not meant to limit the scope of the invention.

[0252] Embodiments herein are further illustrated by the following examples and detailed protocols. However, the examples are merely intended to illustrate embodiments and are not to be construed to limit the scope herein. The contents of all references and published patents and patent applications cited throughout this application are hereby incorporated by reference.

5 [0253] *Example 1. Selective elimination of Leber's hereditary optic neuropathy (LHON)-associated mitochondria! DNA mutations in patient-derived induced pluripotent stem cells using a novel mitochondria-targeted CRISPR system*

[0254] **Significance:** Mitochondria are essential organelles that generate the bulk of cellular energy in the form of ATP from the oxidation of carbohydrates and fats. To carry out this central
10 role in bioenergetics, mitochondria require their own genome, a 16.6 kilobase (kb) circular double-stranded molecule that encodes 37 genes. Each human cell carries hundreds to thousands of copies of mitochondria DNA (mtDNA). In mitochondrial encephalomyopathies, cells typically contain a mixture of both pathogenic and normal mtDNA molecules, a state termed heteroplasmy. There are over 600 known mtDNA mutations associated such mtDNA diseases, which have diverse clinical
15 features, including maternal inheritance (because mtDNA is inherited strictly from the mother), defects in the central and peripheral nervous systems, muscle defects, and exercise intolerance. Due to the inability to transform mitochondrial DNA, there are no approved clinical therapies for the treatment of mitochondrial diseases.

[0255] Leber's hereditary optic neuropathy (LHON) is one of the most common mitochondrial
20 diseases that manifests in the eye. Greater than 95% of LHON patients are characterized by one of three mtDNA point mutations, although dozens of other mtDNA mutations have also been linked to this disease. LHON causes acute bilateral central vision loss affecting preferentially males between the second and third decades of life. Visual loss is often permanent with poor prognosis. Although the pathogenic mtDNA mutations associated with LHON were first identified in complex I subunits
25 more than four decades ago, the pathologic mechanisms of LHON remain incomplete as a result of limitations in cellular and animal models. Most *in vitro* models of LHON utilize trans-mitochondrial cytoplasmic hybrid, or 'cybrids', which are derived from an immortalized osteosarcoma cell line to model the sensitivity of neurons towards mitochondrial bioenergetic defects. Thus, results from cybrid studies have not correlated well with the clinical severity of LHON patients. This may be in

part due to the predominant reliance on the glycolytic pathway by tumor cells for ATP production and thereby circumventing oxidative phosphorylation deficiencies due to complex I mutations.

[0256] Because LHON and other mtDNA diseases begin from heteroplasmic mutations, proof-of-principle attempts have been made to remove the pathogenic mtDNA, thereby shifting the ratio of normal to pathogenic mtDNA. It has been shown that modified transcription activator-like effector nucleases (TALENs) and zinc finger nucleases (ZFNs) that localize into mitochondrion can shift mtDNA heteroplasmy to favor normal mtDNA. Because a high mutational load is required for respiratory dysfunction, a shift in heteroplasmy can restore mitochondrial function to cells carrying mtDNA mutations. While other gene editing methods TALENs and ZFNs are useful biological tools for gene manipulation, each requires protein engineering that can be tedious and inefficient. In contrast, the RNA-guided CRISPR/Cas9 gene editing system requires the design of a short gene-specific single guide RNA (sgRNA) and is amenable for targeting multiple genes simultaneously by multiplexing sgRNAs. In this study we have engineered the CRISPR RNA and Cas9 protein components for trafficking into the mitochondria matrix in order to induce genome editing of mitochondrial DNA. This mtDNA-specific genome editing system could be programmed to target any mutant mtDNA target sequence and could offer new strategy for treating diseases caused by mtDNA mutations.

[0257] Accordingly, the system described herein offers a novel approach for correcting mtDNA disorders, for which there are no effective treatments. In particular, there are provided reagents and method for engineering the CRISPR gene editing technology for mitochondrial localization and selective targeting of LHON-associated mtDNA mutations as a therapy for disease (e.g., mtDNA disease).

[0258] The microbial RNA-mediated adaptive immune system CRISPR, clustered regularly interspaced short palindromic repeats, has become a powerful gene editing tool in eukaryotic cells since the discovery of its mechanism[36]. There are five classes of CRISPR-Cas system, but Class II CRISPR-Cas9 is most widely used because it requires only one endonuclease protein, Cas9, to generate double-strand breaks in the DNA. The Cas9 protein utilizes a synthetic short guide RNA (sgRNA) to direct its DNA cleavage. The sgRNA is composed of a spacer region that is complementary to the DNA target, a CRISPR RNA and a trans-CRISPR RNA duplex. The target

specificity is generated from the spacer domain, and the site of cleavage is determined by a proto-spacer adjacent motif (PAM) upstream of the PAM site.

[0259] Innovation. There are provided reagents and methods useful for engineering the CRISPR gene editing technology for mitochondrial localization and selective targeting of LHON-associated mtDNA mutations as a potential therapy for mitochondrial disease. By utilizing the CRISPR/Cas9 system to target mtDNA, our approach offers four distinct advantages over other similar mtDNA endonuclease technologies. First, the mitoCRISPR system can be easily tailored for any particular mtDNA mutation. Although LHON is primarily caused by three pathogenic mtDNA mutations, there are more than 600 mtDNA mutations linked to mitochondrial diseases according to current records from MITOMAP [43]. Targeting any of these mutations with mitoCRISPR would require the same mitoCas9 enzyme with a unique mito-sgRNA that is complementary to the mtDNA mutation sequence. This is much more versatile than comparable approaches using mitoTALENs or mitoZFNs, which each require two protein arms that must be uniquely engineered to recognize any particular mtDNA mutation. Second, in the design of mitoTALENs and mitoZFNs, one arm is designed to specifically bind to the mutant mtDNA sequence while the other arm must bind to the adjacent wild-type sequence; thus, one protein arm will always bind to both mutant and wild-type mtDNA genomes in a heteroplasmic environment, potentially leading to the steric inhibition of wild-type mtDNA transcription and replication. In contrast, because mitoCRISPR requires only a single mito-sgRNA to target the mutant mtDNA sequence, there should be minimal interference to wild-type mtDNA genomes. Third, while mitoZFNs and their larger mitoTALEN counterparts require two adeno-associated virus (AAV) for expression (one vector for each arm), the mitoCRISPR system can be expressed from a single vector. Fourth, mitoCRISPR offers the ability to simultaneously target multiple mtDNA mutations by multiplexing sgRNA expression. This strategy benefits LHON patients carrying primary and secondary mutations. A similar approach with mitoTALENs or mitoZFNs would demand four mitoTALEN or mitoZFN proteins (two arms per target site), which would require extensive protein engineering and multiple expression vectors.

[0260] Accordingly, one goal is to engineer the CRISPR technology for targeting heteroplasmic mtDNA mutations in LHON-derived iPSC as a means to establish a robust cellular model for therapeutically reversing this disease. Genome editing technologies such as CRISPR, TALENs, and ZFNs, enable the treatment of genetic or acquired diseases by altering defective or pathogenic genes.

However, there are several challenges that hinder the clinical translation of these technologies for nuclear gene therapy. These hurdles include the ability to avoid off target effects, the poor targeting of heterochromatic DNA, the extreme inefficiencies of template-directed homologous DNA recombination (HDR) for gene replacement, and the random nature of non-homologous end joining (NHEJ) for inducing targeted mutagenesis. In contrast to nuclear genome editing, our strategy for targeting mtDNA avoids all of these challenges. First, mtDNA is not wrapped in chromatin and is thus more accessible to targeted endonucleases than most nuclear DNA. Second, at less than 16 kb long, the mitochondria genome is approximately 200,000 times shorter than the nuclear genome, so there is minimal risk of off-target mtDNA cleavage. Third, because mitoCas9 localizes efficiently to the mitochondria with no detectable nuclear localization (Fig. 1B), it is unlikely that any off-target genome editing will occur in the nucleus. Finally, unlike nuclear gene editing applications, which rely on NHEJ for site-directed mutagenesis or HDR for gene replacement therapy, mitoCRISPR is predicated on the selective cleavage and destruction of the mutant mtDNA target sequence. Degradation of double-stranded breaks in mtDNA is the predominant pathway for purging mutations due to the limited repair mechanisms in mitochondria. Thus, the application of mitoCRISPR gene editing may circumvent many of the challenges associated with nuclear gene editing.

[0261] In embodiments, there are provided reagents and methods useful to adapt CRISPR gene editing towards mitochondria: We have successfully targeted the *S. pyogenes* Cas9 protein into the mitochondrial matrix by removing the nuclear localization signals and adding an N-terminal mitochondrial targeting sequence derived from the cytochrome c oxidase subunit VIII (Cox8). We have generated a lentiviral vector with mitoCas9-Flag protein driven by the SFFV promoter and a U6 RNA polymerase III promoter that drives expression of the mitochondrial-localizing sgRNA (mito-sgRNA) (Fig. 1A). Confocal analysis shows highly co-localized signal between mitoCas9-Flag protein with mitochondrial outer membrane marker Tom20 (Fig. 1C). Mitochondrial fractionation also demonstrates an enrichment of Cas9-flag in the mitochondrial fraction (Fig. 1B) relative to the nucleus or cytosol. Figures 5A and 5B further demonstrate that engineered mitoCas9 is successfully localized at mitochondrial matrix according to the Mander's coefficient for the degree of co-localization of mitoCas9 with mitochondria (see Table 7 below).

Table 7

M1 (fraction of Flag signal co-localized with mitochondria): 0.8 ± 0.02 (mean $\pm$ SEM), n= 15 images
--

M2 (fraction of mitochondria with Flag signal): 0.64 ± 0.03 (mean $\pm$ SEM), n= 15 images
--

[0262] In embodiments, there are provided reagents and methods useful to create sgRNAs that efficiently localize into the mitochondrial matrix: We are developing strategies for directing the mitochondrial import of the sgRNA using rationale design. Mechanisms for RNA import in mammalian mitochondria have not been well characterized, but several small RNAs have been isolated from the mitochondrial transcriptome, including the 5S ribosomal RNA, RNaseP and MRP RNA components (reviewed in [46]). Importantly, the delivery of exogenous mRNA or tRNA into mitochondria has been demonstrated by attaching these hairpin loops to the exogenous RNA [47, 48]. Additionally, a modified γ domain of 5S rRNA exhibits enhanced import of RNA into the mitochondria [49]. We will verify the efficacy of these mitochondrial localization loops (mitoloop) for mitochondrial import of the sgRNA (Fig. 2A). We have designed insertion sites of various mitoloops at positions in the sgRNA that have minimal interactions with Cas9, as seen in Fig. 2B. In particular, we have selected the tetraloop and the stem loop 2 to be replaced by mitoloops as mutations in these regions bear no impact on Cas9 binding or activity [50]. Fluorescently labeled sgRNA with various mitoloops have been transfected into mammalian cells. We observe a small degree of colocalization between 488-labeled mito-sgRNA chimera and the mitochondrial marker Tom20 (FIG. 3A). However, with the low efficiency of co-localization by lipofectamine transfection, we have employed a new strategy using lentiviral delivery of the mito-sgRNA driven by a U6 promoter, similar to the construct in Fig. 1A. With constitutive over-expression of the mito-sgRNA chimera, we anticipate greater import of the sgRNA into mitochondria to allow for sufficient mitoCas9 cleavage activity. We have constructed stable cell lines of all the lentiviral expression vectors carrying various permutations of the mito-sgRNA. We will be comparing the efficiency of sgRNA import for each mitoloop using quantitative reverse-transcription PCR (qRT-PCR) of mitoplasts. Mitochondrial outer member is selectively dissolved to form mitoplast as a means to reduce cytoplasmic RNA contamination in our assay.

[0263] In embodiments, there are provided reagents and methods to evaluate and optimize mito-sgRNA specificity for mutant mtDNA target: There are over 600 mutations in mtDNA associated

with devastating encephalomyopathies, and most of these mutations are single point mutations. We are optimizing the specificity of mitoCRISPR mtDNA editing by examining the minimum length of sgRNA to maximize on-target DNA recognition. Truncation of the complementarity region of the sgRNA to 17-18 nt can significantly reduce off-target cleavage with minimal effects on on-target recognition [51]. To examine the specificity of mitoCRISPR towards mtDNA targets, we will develop a nuclear assay with dual chemiluminescence to assess cleavage activity at the wild-type or mutant mtDNA target in the nucleus with nuclear Cas9. Small regions of the mtDNA with wild-type sequence or a single point mutation will be cloned in frame with a chemiluminescence reporter plasmid, which will be transfected into Hela cells. Cleavage of the target should result in insertions/deletions (indels) that disrupt the downstream translation of the reporter or degradation of the linearized reporter plasmid by cellular exonucleases, both of which will result in a loss of signal. Various sgRNA lengths will be tested for DNA recognition specificity. Furthermore, the PAM domain is a critical region in Cas9-mediated DNA melting and RNA-DNA heteroduplex formation. We will test several Cas9 mutants with expanded repertoire of PAM recognition sites and determine if utilizing mutations that create a new PAM domain will enhance the specificity of CRISPR technology for targeting single point mutations in mtDNA.

[0264] In embodiments, there are provided reagents and method to restore mitochondrial function in LHON iPSC with mitoCRISPR: After we have determined the mitoloop design with highest import efficiency and the sgRNA with the highest specificity, we will apply the mitoCRISPR technology to patient-derived iPSC. Rescue of metabolic defects will be assessed using similar assays used to characterize the parental fibroblast and reprogrammed iPSC. It would include measuring changes in heteroplasmy levels, mtDNA copy number normalized to nuclear content, oxygen consumption, ATP levels, complex I activity, mitochondrial number and morphology. Improvements in cell growth and ROS levels will also be determined. This study is the first demonstration of mitoCRISPR as a gene therapy tool for mtDNA disorders and may serve as a therapeutic approach for stem cell transplantation.

[0265] Experimental.

[0266] Creation of Lentiviral Construct (plasmid map will be attached). Addgene vector 57826 was modified by removing both the 5' and 3' nuclear localization signal (NLS) of Cas9 protein. The cytochrome *c* oxidase subunit 8 (COX8) signal was inserted in frame with Cas9 protein

at the 5' terminus as a mitochondrial targeting signal. A 3X flag tag was added in the 3' end of Cas9 as a marker for protein localization. There is a U6 promoter within the vector for expression of guide RNA. We replaced the 2kb dummy sequence downstream of the U6 promoter with the mito-loop sgRNA constructs. The vector is renamed pL_mitoCRISPR.

5 **[0267] Therapy implications.** The system described herein can also utilize the mitoloop to deliver a normal copy of tRNA into mitochondria for functional rescue. This approach/system is most applicable to mtDNA diseases such MELAS (myopathy, encephalopathy, lactic acidosis, stroke-like episodes), MERRF (myoclonic epilepsy, myopathy), and MIDD (maternally inherited diabetes and deafness). However, we can also potentially rescue mtDNA diseases by delivering
10 functional mRNA, codon-optimized for mitochondrial translation, to the mitochondria via the mitoloop. These strategies are tested in available cybrid models. Cellular function and mitochondrial respiration are quantified to determine therapeutic efficacy and efficiency of mitoloop delivery.

[0268] Cpf1. With the recent discovery of Cpf1 as another CRISPR system in Class II, type V,
15 the system described herein also targets Cpf1 into the mitochondria using a similar strategy as Cas9 by fusing a COX8 signal to the 5' of the Cpf1 protein. We utilize the advantages of a shorter guide RNA and ability to modify the 5' end of the guide RNA to test tRNA-mediated delivery of the sgRNA. There have been several reports that nuclear tRNA Glutamine can be imported into the mitochondria. We plan to fuse previously reported tRNA glutamine with CUG and UUG anticodon
20 to the 5' of the Cpf1 sgRNA and determine efficiency of sgRNA delivery into mitochondria. The sequences of the tRNA have been listed in the attached Excel file.

[0269] Development of qPCR strategy for detecting sgRNA. From predicted folding algorithms (e.g., M-Fold and NuPack) and our own empirical results, we have observed the single-stranded sgRNA is highly structured, with particularly strong thermodynamic stability of the sgRNA
25 Stem-loop 2 and Stem-loop 3. Thus, when we tested RT primers which are complementary or downstream of these stem-loop regions, we observed a partial or complete inhibition of the reverse transcription reaction, owing to the unfavorable thermodynamic stability of the folded sgRNA. Thus, we have designed a new RT primer that contains a 12-nt 3'-end sequence which is complementary to a single-stranded region of the sgRNA between Stem-loops 1 and 2 (Figure
30 5). Due to the short length of the resulting complementary DNA (cDNA), which limits our ability to

detect it by quantitative polymerase chain reaction (qPCR), we extended the length of the resulting cDNA by adding a 44-nt stem-loop sequence to the 5'-end of the RT primer. Thus, the complete RT primer is 56-nt long. To detect the PCR amplicon for quantification, we designed a TaqMan probe that spans regions unique to the sgRNA sequence and the RT primer stem-loop. Without this design, we would be unable to detect the sgRNA by qRT-PCR, which would preclude us from confirming its localization to mitochondria.

Table 8. pL_mitoCRISPR plasmid map, Genbank file format. See FIG. 1D.

```

LOCUS       Exported                      11948 bp ds-DNA      circular SYN 13-SEP-2015
DEFINITION   synthetic circular DNA
ACCESSION    .
VERSION      .
KEYWORDS     pL_mitoCRISPR
SOURCE       synthetic DNA construct
  ORGANISM   recombinant plasmid
15  REFERENCE 1 (bases 1 to 11948)
  AUTHORS   Trial User
  TITLE     Direct Submission
  JOURNAL    Exported Jan 25, 2016 from SnapGene 3.0.3
             http://www.snapgene.com
20  FEATURES   Location/Qualifiers
     source    1..11948
               /organism="recombinant plasmid"
               /mol_type="other DNA"
     promoter  5..231
               /note="RSV promoter"
               /note="Rous sarcoma virus enhancer/promoter"
25  LTR        232..412
               /note="5' LTR (truncated)"
               /note="truncated 5' long terminal repeat (LTR) from HIV-1"
30  misc_feature 459..584
               /note="HIV-1 Psi"
               /note="packaging signal of human immunodeficiency virus
               type 1"
     misc_feature 1077..1310
               /note="RRE"
               /note="The Rev response element (RRE) of HIV-1 allows for
               Rev-dependent mRNA export from the nucleus to the
               cytoplasm."
35  promoter  1717..1965
               /note="U6 promoter"
               /note="RNA polymerase III promoter for human U6 snRNA"
     misc_feature 1962..2103
     misc_feature 1962..1991
               /label="sgRNA Target Spacer"
               /note="sgRNA Target Spacer"
45  misc_feature 1967..1972
               /label="BsmBI"
               /note="BsmBI"
               /note="Name: BsmBI;Pattern: cgtctc;Number_of_matches:"

```

2;cut_0_on_positive_strand:
1961^1962;cut_0_on_negative_strand: 1965^1966;inhibited_by:
3': N4-methylcytosine 5': 5-methylcytosine;site_type:
other;restriction site"

5 misc_feature 1981..1986
/label=BsmBI
/note="BsmBI"
/note="Name: BsmBI;Pattern: cgtctc;Number_of_matches:
2;cut_0_on_positive_strand:
10 3846^3847;cut_0_on_negative_strand: 3850^3851;inhibited_by:
3': N4-methylcytosine 5': 5-methylcytosine;site_type:
other;restriction site"

misc_signal 1988..2063
/label=CRISPR RNA
15 /note="CRISPR RNA"

misc_feature 2064..2099
/label=Mitochondrial Localization Loop Spacer
/note="Mitochondrial Localization Loop Spacer"

misc_feature complement(2072..2078)
20 /label=AarI
/note="AarI"

misc_feature 2085..2091
/label=AarI
/note="AarI"

25 misc_feature 2144..2261
/note="cPPT/CTS"
/note="central polypurine tract and central termination
sequence of HIV-1"

misc_feature 2732..2737
30 /note="kozak"

misc_signal 2738..2824
/note="Cox8"

CDS 2831..6931
/codon_start=1
35 /product="Cas9 (Csn1) endonuclease from the Streptococcus
pyogenes Type II CRISPR/Cas system"
/note="Cas9"
/note="generates RNA-guided double strand breaks in DNA"
/translation="DKKYSIGLDIGTNSVGAVITDEYKVPSKKFKVLGNTDRHSIKKN
40 LIGALLFDSGETAEATRLKRTARRRYTRKNNRICYLQEIFSNEMAKVDDSFHRLSEESF
LVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLV DSTDKADLR LIYLALAHMIKF
RGHFLIEGDLNPDNSDVKLF IQLVQTYNQLFEENPINASGVDAKAILSARLSKSRRL
NLIAQLPGEKKNGLFGNLI ALSGLTPNFKSNFDLAEDAKLQLSKDTYDDDLNLLAQI
GDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQ
45 QLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLR
KQRTFDNGSIPHQIHLGELHAILRRQEDFYFPLKDNREKIEKILTFRIPIYVGP LARGN
SRFAWMTRKSEETITPWNFEVVVDKGASAQSFIERM TNFDKNLPNEKVL PKHSLLYEYF
TVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDS
VEISGVEDRFNASLGTYHDL LKIIKDKDFLDNEENEDILEDIVLTLT LFEDREMI EERL
50 KTYAHLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNF MQ
LIHDDSLTFKEDIQKAQVSGQGDSLHEHIANLAGSPA IKKGILQTVKVVDLVKVMGRH
KPENIVIE MARENQT TQKGQKNSRERMKRIE EGIKELGSQILKEHPVENTQLQNEKLYL
YYLQNGRDMYVDQELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPS
EEVVKMKMKNYWRQLLNAKLITQRKFDNLTKAERGGLSELDKAGFIKQLVETRQITKHV
55 AQILDSRMNTKYDENDKLIREVKVITLKS KLVSDFRKDFQFYKVREINNYHHAHDAYLN
AVVGTALIKKYPKLESEFVYGDYKVYDVRKMI AKSEQEIGKATAKYFFYSNIMNFFKTE

ITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKVLSPQVNIIVKKTEVQTGGFSKE
SILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAKEKGKSKKLKSVKELLGIT
IMERSSSFENPIDFLEAKGYKEVKKDLIIKLPKYSLFELNGRKRMLASAGELQKGNEL
ALPSKYVNFYLYASHYEKLKGSPEQNEQQLFVEQHKHYLDEIIIEQISEFSKRVLADA
5 NLDKVL SAYNKHDKPIREQAENIIHLFTLTNLGAPAAFKYFDDTTIDRKRYTSTKEVLD
ATLIHQSIITGLYETRIDLSQLGGD" (SEQ ID NO: 24)

misc_feature 6939..7010
/note="3XFlag_ETKsite"

10 CDS 7019..7075
/codon_start=1
/product="2A peptide from porcine teschovirus-1
polyprotein"
/note="P2A"

15 /note="Eukaryotic ribosomes fail to insert a peptide bond
between the Gly and Pro residues, yielding separate
polypeptides."
/translation="ATNFSLLKQAGDVEENPGP"

CDS 7076..7789
/codon_start=1

20 /product="monomeric derivative of red fluorescent protein
from Entacmaea quadricolor (Merzlyak et al., 2007)"
/note="TagRFP"
/note="mammalian codon-optimized"
/translation="MVSKGEELIKENMHMKLYMEGTVNNHHFKCTSEGEKPYEGTQTM

25 RIKVVEGGPLPFAFDILATSFMYGSRTFINHTQGIPDFFKQSFPEGFTWERVTTYEDGG
VLTATQDTSLDGGLIYNVKIRGVNFPSNGPVMQKKT LGWEANTEMLYPADGGLEGRSD
MALKLVGGGHLICNFKTTYRSKKPAKNLKMGPVYVDHRLRIKEADKETYVEQHEVAV
ARYCDLPSKLGHKLN"

misc_feature 7805..8393
/note="WPRE"

30 /note="woodchuck hepatitis virus posttranscriptional
regulatory element"

CDS complement(8276..8287)
/codon_start=1

35 /product="Factor Xa recognition and cleavage site"
/note="Factor Xa site"
/translation="IEGR"

LTR 8465..8698
/note="3' LTR (Delta-U3)"

40 /note="self-inactivating 3' long terminal repeat (LTR) from
HIV-1"

polyA_signal 8776..8897
/note="SV40 poly(A) signal"
/note="SV40 polyadenylation signal"

45 rep_origin 8937..9072
/note="SV40 ori"

promoter /note="SV40 origin of replication"
complement(9093..9111)
/note="T7 promoter"

50 primer_bind /note="promoter for bacteriophage T7 RNA polymerase"
complement(9121..9137)
/note="M13 fwd"

55 /note="common sequencing primer, one of multiple similar
variants"

rep_origin 9279..9734
/direction=RIGHT

```

5         promoter      /note="f1 ori"
                        /note="f1 bacteriophage origin of replication; arrow
                        indicates direction of (+) strand synthesis"
                        9760..9864
                        /gene="bla"
                        /note="AmpR promoter"
10        CDS           9865..10725
                        /codon_start=1
                        /gene="bla"
                        /product="beta-lactamase"
                        /note="AmpR"
                        /note="confers resistance to ampicillin, carbenicillin, and
                        related antibiotics"
                        /translation="MSIQHFRVALIPFFAAFCCLPVFAHPETLVKVKDAEDQLGARVGYI
15        ELDLNSGKILESFRPEERFPMSTFKVLLCGAVLSRIDAGQEQLGRRRIHYSQNDLVEYS
                        PVTEKHLTDGMTVRELCSAAITMSDNTAANLLLTIGGPKELTAFLHNMGDHVTRLDRW
                        EPENEAIPNDERDTTMPVAMATTLRKLLTGELLTLASRQQILIDWMEADKVAGPLLRSA
                        LPAGWFIADKSGAGERGSRGIIAALGPDGKPSRIVVIYTTGSQATMDERNRQIAEIGAS
                        LIKHW"
20        rep_origin    10896..11484
                        /direction=RIGHT
                        /note="ori"
                        /note="high-copy-number ColE1/pMB1/pBR322/pUC origin of
                        replication"
25        protein_bind  11772..11793
                        /bound_moiety="E. coli catabolite activator protein"
                        /note="CAP binding site"
                        /note="CAP binding activates transcription in the presence
                        of cAMP."
30        promoter      11808..11838
                        /note="lac promoter"
                        /note="promoter for the E. coli lac operon"
                        protein_bind  11846..11862
                        /bound_moiety="lac repressor encoded by lacI"
35        /note="lac operator"
                        /note="The lac repressor binds to the lac operator to
                        inhibit transcription in E. coli. This inhibition can be
                        relieved by adding lactose or
                        isopropyl-beta-D-thiogalactopyranoside (IPTG)."
```

40 primer_bind 11870..11886

```

                        /note="M13 rev"
                        /note="common sequencing primer, one of multiple similar
                        variants"
45        promoter      11907..11925
                        /note="T3 promoter"
                        /note="promoter for bacteriophage T3 RNA polymerase"
ORIGIN
1   ttaatgtagt cttatgcaat actcctgtag tcttgcaaca tggtaacgat gagttagcaa
61  catgccttac aaggagagaa aaagcaccgt gcatgccgat tggtggaagt aaggtggtac
50  121 gatcgtgcct tattaggaag gcaacagacg ggtctgacat ggattggacg aaccactgaa
181 ttgccgcatt gcagagatat tgtatttaag tgcctagctc gatacataaa cgggtctctc
241 tggttagacc agatctgagc ctgggagctc tctggctaac tagggaaccc actgcttaag
301 cctcaataaa gcttgccctt agtgcttcaa gtagtggttg cccgtctggt gtgtgactct
361 ggtaactaga gatccctcag acccttttag tcagtgtgga aaatctctag cagtggcgcc
55  421 cgaacaggga cttgaaagcg aaagggaaac cagaggagct ctctcgacgc aggactcggc
481 ttgctgaagc gcgcacggca agaggcgagg ggcggcgact ggtgagtacg ccaaaaattt

```

541	tgactagcgg	aggctagaag	gagagagatg	ggtgcgagag	cgtcagtatt	aagcggggga
601	gaattagatc	gcatgaggaa	aaaattcggg	taaggccagg	gggaaagaaa	aaatataaat
661	taaaacatat	agtatgggca	agcaggggagc	tagaacgatt	cgcagttaat	cctggcctgt
721	tagaaacatc	agaaggctgt	agacaaatac	tgggacagct	acaaccatcc	cttcagacag
5	781	gatcagaaga	acttagatca	ttatataata	cagtagcaac	cctctattgt
841	ggatagagat	aaaagacacc	aaggaagcct	tagacaagat	agaggaagag	caaaacaaaa
901	gtaagaccac	cgcacagcaa	gcggccgctg	atcttcagac	ctggaggagg	agatatgagg
961	gacaattgga	gaagtgaatt	atataaatat	aaagtagtaa	aaattgaacc	attaggagta
1021	gcacccacca	aggcaaagag	aagagtgggt	cagagagaaa	aaagagcagt	gggaatagga
10	1081	gctttgttcc	ttgggttctt	gggagcagca	ggaagcacta	tgggcgcagc
1141	ctgacggtac	aggccagaca	attattgtct	ggtatagtgc	agcagcagaa	caatttgctg
1201	agggctattg	aggcgcaaca	gcatctgttg	caactcacag	tctggggcat	caagcagctc
1261	caggcaagaa	tcctggctgt	ggaaagatac	ctaaaggatc	aacagctcct	ggggatttgg
1321	ggttgctctg	gaaaactcat	ttgcaccact	gctgtgcctt	ggaatgctag	ttggagtaat
15	1381	aaatctctgg	aacagatttg	gaatcacacg	acctggatgg	agtgggacag
1441	aattacacaa	gcttaataca	ctccttaatt	gaagaatcgc	aaaaccagca	agaaaagaat
1501	gaacaagaat	tattggaatt	agataaatgg	gcaagtttgt	ggaattgggt	taacataaca
1561	aattggctgt	ggtatataaa	attattcata	atgatagtag	gaggcttggg	aggtttaaga
1621	atagtttttg	ctgtactttc	tatagtgaat	agagttaggc	agggatatcc	accattatcg
20	1681	tttcagaccc	acctcccaac	cccagaggga	cccagagagg	gcctatttcc
1741	tcataatttg	atatacgata	caaggctgtt	agagagataa	ttagaattaa	tttgactgta
1801	aacacaaaaga	tattagtaca	aaatacgtga	cgtagaaagt	aataatttct	tgggtagttt
1861	gcagttttta	aattatgttt	taaaatggac	tatcatatgc	ttaccgtaac	ttgaaagtat
1921	ttcgattttct	tggctttata	tatcttgtgg	aaaggacgaa	acaccggaga	cggcgggccgc
25	1981	cgtctctgtt	ttagagctag	aaatagcaag	ttaaaataag	gctagtccgt
2041	aaaaagtggc	accgagtcgg	tgccaccggc	cgcaggtggg	atcccacctg	ccatggtttt
2101	ttgaattcta	gatcttgaga	caaatggcag	tattcatcca	caatttttaa	agaaaagggg
2161	ggattggggg	gtacagtgca	ggggaaagaa	tagtagacat	aatagcaaca	gacatacaaa
2221	ctaaagaatt	acaaaaacaa	attacaaaaa	ttcaaaaatt	tcgggtttat	tacagggaca
30	2281	gcagagatcc	actttggcgc	cggctcgagc	gagctgcagt	aacgccattt
2341	tgaaaaatag	caaaaccaag	atagagaagt	tcagatcaag	ggcgggtaca	tgaaaatagc
2401	taacgttggg	ccaaacagga	tatctgcggg	gagcagtttc	ggccccggcc	cggggccaag
2461	aacagatggt	caccgcagtt	tcggccccgg	cccagggcca	agaacagatg	gtccccagat
2521	atggcccaac	cctcagcagt	ttcttaagac	ccatcagatg	tttccaggct	cccccaagga
35	2581	cctgaaatga	ccctgcgcct	tatttgaatt	aaccaatcag	cctgcttctc
2641	gcgcgcttct	gcttcccgag	ctctataaaa	gagctcacia	cccctcactc	ggcgcgccag
2701	tcctccgaca	gactgagtcg	gatcaactag	tgccaccatg	tcgctcctga	cgccgctgct
2761	gctgcggggc	ttgacaggct	cggccccggc	gctcccagtg	ccgcgcgcca	agatccattc
2821	ggttgatccg	gacaagaagt	acagcatcgg	cctggacatc	ggcaccaact	ctgtgggctg
40	2881	ggccgtgatc	accgacgagt	acaaggtgcc	cagcaagaaa	ttcaaggtgc
2941	cgaccggcac	agcatcaaga	agaacctgat	cggagccctg	ctgttcgaca	gcggcgaaac
3001	agccgaggcc	acccggctga	agagaaccgc	cagaagaaga	tacaccagac	ggaagaaccg
3061	gatctgctat	ctgcaagaga	tcttcagcaa	cgagatggcc	aaggtggacg	acagcttctt
3121	ccacagactg	gaagagtcct	tcctggtgga	agaggataag	aagcacgagc	ggcaccat
45	3181	cttcggcaac	atcgtggacg	aggtggccta	ccacgagaag	tacccacca
3241	gagaaagaaa	ctggtggaca	gcaccgacaa	ggccgacctg	cggctgatct	atctggccct
3301	ggcccacatg	atcaagttcc	ggggccactt	cctgatcgag	ggcgacctga	accccgacaa
3361	cagcgacgtg	gacaagctgt	tcattccagct	ggtgcagacc	tacaaccagc	tggtcgagga
3421	aaaccccatc	aacgccagcg	gcgtggacgc	caaggccatc	ctgtctgcca	gactgagcaa
50	3481	gagcagacgg	ctggaaaatc	tgatgcacca	gctgcccggc	gagaagaaga
3541	cggaacactg	attgccctga	gcctgggcct	gacccccaac	ttcaagagca	acttcgacct
3601	ggccgaggat	gccaaactgc	agctgagcaa	ggacacctac	gacgacgacc	tggaacaact
3661	gctggcccgag	atcggcgacc	agtacgccga	cctgtttctg	gccgccaaga	acctgtccga
3721	cgccatcctg	ctgagcgaca	tcctgagagt	gaacaccgag	atcaccaagg	ccccctgag
55	3781	cgctctatg	atcaagagat	acgacgagca	ccaccaggac	ctgacctgc
3841	cgtgcggcag	cagctgcctg	agaagtacaa	agagattttc	ttcgaccaga	gcaagaacgg

	3901	ctacgccggc	tacattgacg	gcgagagccag	ccaggaagag	ttctacaagt	tcatcaagcc
	3961	catcctggaa	aagatggacg	gcaccgagga	actgctcgtg	aagctgaaca	gagaggacct
	4021	gctgcggaag	cagcggacct	tcgacaacgg	cagcatcccc	caccagatcc	acctgggaga
5	4081	gctgcacgcc	attctgcggc	ggcaggaaga	tttttaccga	ttcctgaagg	acaaccggga
	4141	aaagatcgag	aagatcctga	ccttcgcgat	cccctactac	gtggggccctc	tggccagggg
	4201	aaacagcaga	ttcgcctgga	tgaccagaaa	gagcgaggaa	accatcacc	cctggaactt
	4261	cgaggaagtg	gtggacaagg	gcgcttcgc	ccagagcttc	atcgagcgga	tgaccaactt
	4321	cgataagaac	ctgcccacg	agaaggtgct	gcccagcac	agcctgctgt	acgagtactt
10	4381	caccgtgtat	aacgagctga	ccaaagtga	atacgtgacc	gagggaatga	gaaagcccgc
	4441	cttcctgagc	ggcgagcaga	aaaaggccat	cgtggacctg	ctgttcaaga	ccaaccggaa
	4501	agtgaccgtg	aagcagctga	aagaggacta	cttcaagaaa	atcgagtgtc	tcgactccgt
	4561	ggaaatctcc	ggcgtggaag	atcgggttcaa	cgccctccctg	ggcacatacc	acgatctgct
	4621	gaaaattatc	aaggacaagg	acttctctgga	caatgaggaa	aacgaggaca	ttctggaaga
	4681	tatcgtgctg	accctgacac	tgtttgagga	cagagagatg	atcgaggaa	ggctgaaaac
15	4741	ctatgcccac	ctgttcgacg	acaaagtgat	gaagcagctg	aagcggcgga	gatacaccgg
	4801	ctggggcagg	ctgagccgga	agctgatcaa	cggcatccgg	gacaagcagt	ccggcaagac
	4861	aatcctggat	ttcctgaagt	ccgacggctt	cgccaacaga	aacttcatgc	agctgatcca
	4921	cgacgacagc	ctgaccttta	aagaggacat	ccagaaaagg	caggtgtccg	gccagggcga
20	4981	tagcctgcac	gagcacattg	ccaatctggc	cggcagcccc	gccattaaga	agggcatcct
	5041	gcagacagtg	aaggtggtgg	acgagctcgt	gaaagtgatg	ggccggcaca	agcccgagaa
	5101	catcgtgatc	gaaatggcca	gagagaacca	gaccacccag	aagggaacaga	agaacagccg
	5161	cgagagaatg	aagcggatcg	aagagggcat	caaagagctg	ggcagccaga	tcctgaaaga
	5221	acaccccgtg	gaaaacaccc	agctgcagaa	cgagaagctg	tacctgtact	acctgcagaa
25	5281	tggggcggat	atgtacgtgg	accaggaact	ggacatcaac	cggctgtccg	actacgatgt
	5341	ggaccatata	gtgcctcaga	gctttctgaa	ggacgactcc	atcgacaaca	aggtgtgtgac
	5401	cagaagcgac	aagaaccggg	gcaagagcga	caacgtgccc	tcggaagagg	tcgtgaagaa
	5461	gatgaagaac	tactggcggc	agctgctgaa	cgccaagctg	attaccacaga	gaaagttcga
	5521	caatctgacc	aaggccgaga	gaggcggcct	gagcgaactg	gataaggccg	gcttcatcaa
30	5581	gagacagctg	gtggaaaacc	ggcagatcac	aaagcacgtg	gcacagatcc	tggactcccg
	5641	gatgaacata	aagtacgacg	agaattgaca	gctgatccgg	gaagtgaag	tgatcacctc
	5701	gaagtccaag	ctggtgtccg	atttccggaa	ggatttccag	ttttacaaag	tcgcgcgagat
	5761	caacaactac	caccacgccc	acgacgccta	cctgaacgcc	gtcgtgggaa	ccgcccgtgat
	5821	caaaaagtac	cctaagctgg	aaagcaggtt	cgtgtacggc	gactacaagg	tgtacgacgt
35	5881	gcggaagatg	atcgccaaga	gcgagcagga	aatcggcaag	gctaccgcca	agtacttctt
	5941	ctacagcaac	atcatgaact	ttttcaagac	cgagattacc	ctggccaacg	gcgagatccg
	6001	gaagcggcct	ctgatcgaga	caaacggcga	aaccggggag	atcgtgtggg	ataagggccg
	6061	ggatttttgc	accgtgcgga	aagtgtctgag	catgccccaa	gtgaatatcg	tgaaaaagac
	6121	cgaggtgcag	acaggcggct	tcagcaaaga	gtctatcctg	cccaagagga	acagcgataa
40	6181	gctgatcgcc	agaaagaagg	actgggaccc	taagaagtac	ggcggcttcg	acagccccac
	6241	cgtggcctat	tctgtgctgg	tggtggccaa	agtggaaaag	ggcaagtcca	agaaactgaa
	6301	gagtgtgaaa	gagctgctgg	ggatcaccat	catggaaaaga	agcagcttcg	agaagaatcc
	6361	catcgacttt	ctggaagcca	agggctacaa	agaagtgaag	aaggacctga	tcatcaagct
	6421	gcctaagtac	tcctgttctg	agctggaaaa	cggccgggaag	agaatgctgg	cctctgccgg
45	6481	cgaactgcag	aagggaacg	aactggccct	gccctccaaa	tatgtgaact	tcctgtacct
	6541	ggccagccac	tatgagaagc	tgaagggtct	ccccgaggat	aatgagcaga	aacagctgtt
	6601	tgtggaacag	cacaagcact	acctggacga	gatcatcgag	cagatcagcg	agttctccaa
	6661	gagagtgate	ctggccgacg	ctaactctgga	caaagtgtctg	tcgccttaca	acaagcaccg
	6721	ggataagccc	atcagagagc	aggccgagaa	tatcatccac	ctgtttaccc	tgaccaatct
	6781	gggagccctt	gccgccttca	agtactttga	caccaccatc	gaccggaaga	ggtacaccag
50	6841	caccaaagag	gtgctggacg	ccaccctgat	ccaccagagc	atcacgggcc	tgtacgagac
	6901	acggatcgac	ctgtctcagc	tgggaggcga	cgctagcgac	tataaggacc	acgacgggaga
	6961	ctacaaggat	catgatattg	attacaaaga	cgatgacgat	aagcctagcg	gcagcggcgc
	7021	caccaacttc	agcctgctga	agcaggccgg	cgacgtggag	gagaaccccc	gccccatggg
	7081	gtctaagggc	gaagagctga	ttaaggagaa	catgcacatg	aagctgtata	tggagggcac
55	7141	cgtgaacaac	caccacttca	agtgcacatc	cgaggggcga	ggcaagccct	acgagggcac
	7201	ccagaccatg	agaatcaagg	tggtcgaggg	cggccctctc	cccttcgcct	tcgacatcct

7261 ggctaccagc ttcatgtacg gcagcagaac cttcatcaac cacacccagg gcatccccga
 7321 cttctttaag cagtccttcc ctgagggctt cacatgggag agagtcacca catacgaaga
 7381 cgggggctg ctgaccgcta cccaggacac cagcctccag gacggctgcc tcatctacaa
 7441 cgtcaagatc agaggggtga acttcccatc caacggccct gtgatgcaga agaaaacact
 5 7501 cggctgggag gccaacaccg agatgctgta ccccgctgac ggcggcctgg aaggcagaag
 7561 cgacatggcc ctgaagctcg tgggcggggg ccacctgatc tgcaacttca agaccacata
 7621 cagatccaag aaaccgcta agaacctcaa gatgcccggc gtctactatg tggaccacag
 7681 actggaaga atcaaggagg ccgacaaaga gacctacgtc gagcagcacg aggtggctgt
 10 7741 ggccagatac tgcgacctcc ctagcaaact ggggcacaaa ctttaattgaa cgcgttaagt
 7801 cgacaatcaa cctctggatt acaaaatttg tgaaagattg actggtattc ttaactatgt
 7861 tgctcctttt acgctatgtg gatacgctgc tttaatgcct ttgtatcatg ctattgcttc
 7921 ccgatggctt ttcattttct cctccttgta taaatcctgg ttgctgtctc tttatgagga
 7981 gttgtggccc gttgtcaggg aacgtggcgt ggtgtgcact gtgtttgctg acgcaacccc
 8041 cactggttgg gccattgcca ccacctgtca gctcctttcc gggactttcg ctttccccct
 15 8101 ccctattgac acggcggaac tcatcgccgc ctgccttgcc cgctgctgga caggggctcg
 8161 gctgttgggc actgacaatt ccgtggtggt gtcggggaaa tcatcgctcc ttccttggct
 8221 gctcgctgt gttgccacct ggattctgcg cgggacgtcc ttctgctacg tcccttcggc
 8281 cctcaatcca gcggaccttc cttcccgcg cctgctgcgg gctctgcggc ctcttcggc
 8341 tcttcgcctt cgccctcaga cgagtcggat ctccttttgg gccgctccc cgcgtcgact
 20 8401 ttaagaccaa tgacttaciaa ggcagctgta gatcttagcc actttttaaa agaaaagggg
 8461 ggactggaag ggctaattca ctcccaacga agacaagatc tgctttttgc ttgtactggg
 8521 tctctctggg tagaccagat ctgagcctgg gagctctctg gctaactagg gaacccactg
 8581 cttaagcctc aataaagctt gccttgagtg cttcaagtag tgtgtgccc tctgttgtgt
 25 8641 gactctggta actagagatc cctcagaccc ttttagtcag tgtggaaaat ctctagcagt
 8701 acgtatagta gttcatgtca tcttattatt cagtatttat aacttgcaaa gaaatgaata
 8761 tcagagagtg agaggaaact gtttattgca gcttataatg gttacaaata aagcaatagc
 8821 atcacaaatt tcacaaataa agcatttttt tcaactgcatt ctagttgtgg tttgtccaaa
 8881 ctcatcaatg tatcttatca tgtctggctc tagctatccc gcccttaact ccgcccattc
 8941 cgcccctaac tccgcccagt tccgcccatt ctcgccccca tggctgacta atttttttta
 30 9001 tttatgcaga ggcgagggc gcctcgccct ctgagctatt ccagaagtag tgaggaggct
 9061 tttttggagg cctagggagc taccacattc gccctatagt gactgctatt acgcgcgctc
 9121 actggccgct gttttacaac gtcgtgactg ggaaaaccct ggcgttacc aacttaatcg
 9181 ccttgacgca catccccctt tcgcccagct gcgtaatagc gaagaggccc gcaccgatcg
 9241 cccttcccaa cagttgcgca gcctgaatgg cgaatgggac gcgcccgtga gcggcgctt
 35 9301 aagcgcgcg ggtgtggtgg ttacgcgcag cgtgaccgct acacttgcca gcgcccagc
 9361 gcccgctcct ttcgctttct tcccttcctt tctcgccaag ttcccggtt tcccccgtca
 9421 agctctaaat cgggggctcc ctttaggggt ccgatttagt gctttacggc acctcgaccc
 9481 caaaaaactt gattaggggt atggttcacg tagtgggcca tcgcccgtat agacggtttt
 9541 tcgccccttg acgttgaggt ccacgttctt taatagtggg ctcttggtcc aaactggaac
 40 9601 aacactcaac cctatctcgg tctattcttt tgatttataa gggattttgc cgatttcggc
 9661 ctattgggta aaaaatgagc tgattttaaca aaaatttaac gcgaatttta acaaaatatt
 9721 aacgcttaca atttaggtgg cacttttcgg ggaaatgtgc gcggaacccc tatttggtta
 9781 tttttctaaa tacattcaaa tatgtatccg ctcatgagac aataaccctg ataaatgctt
 45 9841 caataatatt gaaaaaggaa gagtatgagt attcaacatt tccgtgtcgc ccttattccc
 9901 ttttttgcg ctttttgct tccgttttt gctcaccag aaacgctggt gaaagtaaaa
 9961 gatgctgaag atcagttggg tgcacgagtg ggttacatcg aactggatct caacagcggt
 10021 aagatccttg agagttttcg cccgaagaa cgttttccaa tgatgagcac ttttaaagtt
 10081 atgctatgtg gcgcggtatt atcccgatt gacgcggggc aagagcaact cggctcggc
 10141 atacactatt ctcagaatga cttggttag tactaccag tcacagaaaa gcatcttacg
 50 10201 gatggcatga cagtaagaga attatgcagt gctgccataa ccatgagtga taactctg
 10261 gccaaactac ttctgacaac gatcgaggga ccgaaggagc taaccgcttt tttgcacaac
 10321 atgggggatc atgtaactcg ccttgatcgt tgggaacgg agctgaatga agccatacca
 10381 aacgacgagc gtgacaccac gatgcctgta gcaatggcaa caacgttgcg caaactatta
 10441 actggcgaac tacttactct agcttcccg caacaattaa tagactggat ggaggcggt
 55 10501 aaagttgcag gaccacttct gcgctcggcc cttccggctg gctgggttat tgctgataaa
 10561 tctggagccg gtgagcgtgg gtctcgcggt atcattgcag cactggggcc agatggtaag

```

10621 ccctcccgta tcgtagttat ctacacgacg gggagtcagg caactatgga tgaacgaaat
10681 agacagatcg ctgagatagg tgcctcactg attaagcatt ggtaactgtc agaccaagtt
10741 tactcatata tacttttagat tgattttaaaa cttcattttt aattttaaaag gatctagggtg
10801 aagatccttt ttgataatct catgacccaaa atcccttaac gtgagttttc gttccactga
5 10861 gcgtcagacc ccgtagaaaa gatcaaagga tcttcttgag atcctttttt tctgcgcgta
10921 atctgctgct tgcaaacaaa aaaaccaccg ctaccagcgg tggtttggtt gccggatcaa
10981 gagctaccaa ctctttttcc gaaggtaact ggcttcagca gagcgcagat accaaatact
11041 gttcttctag tgtagccgta gttaggccac cacttcaaga actctgtagc accgcctaca
11101 tacctcgctc tgctaatacct gttaccagtg gctgctgccg gtggcgataa gtcgtgtctt
10 11161 accgggttgg actcaagacg atagttaccg gataaggcgc agcggtcggg ctgaacgggg
11221 ggttcgtgca cacagcccag cttggagcga acgacctaca ccgaactgag atacctacag
11281 cgtgagctat gagaaagcgc cagccttccc gaaggagaga aggcggacag gtatccggta
11341 agcggcaggg tcggaacagg agagcgcacg agggagcttc cagggggaaa cgcctgggat
11401 ctttatagtc ctgtcggggt tcgccacctc tgacttgagc gtcgattttt gtgatgctcg
15 11461 tcaggggggc ggagcctatg gaaaaacgcc agcaacgcgg cctttttacg gttcctggcc
11521 ttttgctggc cttttgctca catgttcttt cctgcgttat cccctgattc tgtggataac
11581 cgtattaccg cctttgagtg agctgatacc gctcgccgca gccgaacgac cgagcgcagc
11641 gagtcagtga gcgaggaagc ggaagagcgc ccaatacgca aaccgcctct ccccgcgctg
11701 tggccgattc attaatgcag ctggcacgac aggtttcccg actggaaagc gggcagtgag
20 11761 cgcaacgcaa ttaatgtgag ttagctcact cattagggac cccaggcttt acactttatg
11821 cttccggctc gtatgttggtg tggaattgtg agcggataac aatttcacac aggaaacagc
11881 tatgaccatg attacgcca ggcgcgaatt aaccctcact aaagggaaca aaagctggag
11941 ctgcaagc (SEQ ID NO: 125)
//
25

```

[0270] References (Example 1)

- [0271] [1] Huoponen, K., et al., A new mtDNA mutation associated with Leber hereditary optic neuroretinopathy. American Journal of Human Genetics, 1991. 48(6): p. 1147-53.: [2] Wallace, D.C., et al., Mitochondrial DNA mutation associated with Leber's hereditary optic neuropathy. Science, 1988. 242(4884): p. 1427-30.: [3] Yu-Wai-Man, P., P.G. Griffiths, and P.F. Chinnery, Mitochondrial optic neuropathies - disease mechanisms and therapeutic strategies. Prog Retin Eye Res, 2011. 30(2): p. 81-114.: [4] Pan, B.X., et al., Mathematically modeling the involvement of axons in Leber's hereditary optic neuropathy. Invest Ophthalmol Vis Sci, 2012. 53(12): p. 7608-17.: [5]. oilkonda, R.D. and J. Guy, Leber's Hereditary Optic Neuropathy-Gene Therapy: From Benchtop to Bedside. J Ophthalmol, 2011. 2011: p. 179412.: [6] Erickson, R.P., Leber's optic atrophy, a possible example of maternal inheritance. Am J Hum Genet, 1972. 24(3): p. 348-9.: [7] Carelli, V., F.N. Ross-Cisneros, and A.A. Sadun, Mitochondrial dysfunction as a cause of optic neuropathies. Prog Retin Eye Res, 2004. 23(1): p. 53-89.: [8] Huoponen, K., Leber hereditary optic neuropathy: clinical and molecular genetic findings. Neurogenetics, 2001. 3(3): p. 119-25.: [9] Bu, X.D. and J.I. Rotter, X chromosome-linked and mitochondrial gene control of Leber hereditary optic neuropathy: evidence from segregation analysis for dependence on X chromosome inactivation. Proc

- Natl Acad Sci U S A, 1991. 88(18): p. 8198-202.: [10] Ji, Y., et al., Evaluation of the X-linked modifier loci for Leber hereditary optic neuropathy with the G11778A mutation in Chinese. *Mol Vis*, 2010. 16: p. 416-24.: [11] Hudson, G., et al., Identification of an X-chromosomal locus and haplotype modulating the phenotype of a mitochondrial DNA disorder. *Am J Hum Genet*, 2005. 77(6): p. 1086-91.: [12] Chen, J.D. and M.J. Denton, X-chromosomal gene in Leber hereditary optic neuroretinopathy. *Am J Hum Genet*, 1991. 49(3): p. 692-3.: [13] Oostra, R.J., et al., No evidence for 'skewed' inactivation of the X-chromosome as cause of Leber's hereditary optic neuropathy in female carriers. *Hum Genet*, 1996. 97(4): p. 500-5.: [14] Giordano, C., et al., Oestrogens ameliorate mitochondrial dysfunction in Leber's hereditary optic neuropathy. *Brain*, 2011. 134(Pt 1): p. 220-34.: [15] Pisano, A., et al., Targeting estrogen receptor beta as preventive therapeutic strategy for Leber's hereditary optic neuropathy. *Hum Mol Genet*, 2015. 24(24): p. 6921-31.: [16] Bristow, E.A., et al., The distribution of mitochondrial activity in relation to optic nerve structure. *Arch Ophthalmol*, 2002. 120(6): p. 791-6.: [17] Sadun, A.A., et al., Leber's hereditary optic neuropathy differentially affects smaller axons in the optic nerve. *Trans Am Ophthalmol Soc*, 2000. 98: p. 223-32; discussion 232-5.: [18] King, M.P. and G. Attardi, Human cells lacking mtDNA: repopulation with exogenous mitochondria by complementation. *Science*, 1989. 246(4929): p. 500-3.: [19] Schoeler, S., et al., Glutathione depletion in antioxidant defense of differentiated NT2-LHON cybrids. *Neurobiol Dis*, 2007. 25(3): p. 536-44.: [20] Zanna, C., et al., Caspase-independent death of Leber's hereditary optic neuropathy cybrids is driven by energetic failure and mediated by AIF and Endonuclease G. *Apoptosis*, 2005. 10(5): p. 997-1007.: [21] Baracca, A., et al., Severe impairment of complex I-driven adenosine triphosphate synthesis in leber hereditary optic neuropathy cybrids. *Arch Neurol*, 2005. 62(5): p. 730-6.: [22] Kirches, E., LHON: Mitochondrial Mutations and More. *Current Genomics*, 2011. 12(1): p. 44-54.: [23] Sadun, A.A., et al., Extensive investigation of a large Brazilian pedigree of 11778/haplogroup J Leber hereditary optic neuropathy. *Am J Ophthalmol*, 2003. 136(2): p. 231-8.: [24] Sadun, A.A., et al., Effect of EPI-743 on the clinical course of the mitochondrial disease Leber hereditary optic neuropathy. *Arch Neurol*, 2012. 69(3): p. 331-8.: [25] Meyerson, C., G. Van Stavern, and C. McClelland, Leber hereditary optic neuropathy: current perspectives. *Clin Ophthalmol*, 2015. 9: p. 1165-76.: [26] Koilkonda, R.D., et al., Safety and effects of the vector for the Leber hereditary optic neuropathy gene therapy clinical trial. *JAMA Ophthalmol*, 2014. 132(4): p. 409-20.: [27] Lightowlers, R.N., R.W. Taylor, and D.M. Turnbull, Mutations causing mitochondrial disease: What is new and what challenges remain?

- Science, 2015. 349(6255): p. 1494-9.: [28] Alexeyev, M., et al., The maintenance of mitochondrial DNA integrity--critical analysis and update. Cold Spring Harb Perspect Biol, 2013. 5(5): p. a012641.: [29] Bacman, S.R., et al., Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. Nat Med, 2013. 19(9): p. 1111-3.: [30] Gammage, P.A., et al., Mitochondrially targeted ZFNs for selective degradation of pathogenic mitochondrial genomes bearing large-scale deletions or point mutations. EMBO Mol Med, 2014. 6(4): p. 458-66.: [31] Hashimoto, M., et al., MitoTALEN: A General Approach to Reduce Mutant mtDNA Loads and Restore Oxidative Phosphorylation Function in Mitochondrial Diseases. Mol Ther, 2015. 23(10): p. 1592-9.: [32] Reddy, P., et al., Selective elimination of mitochondrial mutations in the germline by genome editing. Cell, 2015. 161(3): p. 459-69.: [33] Tachibana, M., et al., Mitochondrial gene replacement in primate offspring and embryonic stem cells. Nature, 2009. 461(7262): p. 367-72.: [34] Ma, H., et al., Metabolic rescue in pluripotent cells from patients with mtDNA disease. Nature, 2015. 524(7564): p. 234-8.: [35] Deuse, T., et al., SCNT-derived ESCs with mismatched mitochondria trigger an immune response in allogeneic hosts. Cell Stem Cell, 2015. 16(1): p. 33-8.: [36] Jinek, M., et al., A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science, 2012. 337(6096): p. 816-21.: [37] Vierbuchen, T., et al., Direct conversion of fibroblasts to functional neurons by defined factors. Nature, 2010. 463(7284): p. 1035-41.: [38] Zhang, Y., et al., Rapid single-step induction of functional neurons from human pluripotent stem cells. Neuron, 2013. 78(5): p. 785-98.: [39] Tanaka, T., et al., Generation of retinal ganglion cells with functional axons from human induced pluripotent stem cells. Sci Rep, 2015. 5: p. 8344.: [40] Folmes, C.D., et al., Disease-causing mitochondrial heteroplasmy segregated within induced pluripotent stem cell clones derived from a patient with MELAS. Stem Cells, 2013. 31(7): p. 1298-308.: [41] Schwartz, S.D., et al., Human embryonic stem cell-derived retinal pigment epithelium in patients with age-related macular degeneration and Stargardt's macular dystrophy: follow-up of two open-label phase 1/2 studies. Lancet, 2015. 385(9967): p. 509-16.: [42] Song, W.K., et al., Treatment of macular degeneration using embryonic stem cell-derived retinal pigment epithelium: preliminary results in Asian patients. Stem Cell Reports, 2015. 4(5): p. 860-72.: [43] Brandon, M.C., et al., MITOMAP: a human mitochondrial genome database--2004 update. Nucleic Acids Res, 2005. 33(Database issue): p. D611-3.: [44] Howell, N., et al., A heteroplasmic LHON family: tissue distribution and transmission of the 11778 mutation. Am J Hum Genet, 1994. 55(1): p. 203-6.: [45] Phillips, N.R., M.L. Sprouse, and R.K. Roby, Simultaneous quantification of mitochondrial DNA

copy number and deletion ratio: a multiplex real-time PCR assay. *Sci Rep*, 2014. 4: p. 3887.: [46] Schneider, A., Mitochondrial tRNA import and its consequences for mitochondrial translation. *Annu Rev Biochem*, 2011. 80: p. 1033-53.: [47] Wang, G., et al., PNPASE regulates RNA import into mitochondria. *Cell*, 2010. 142(3): p. 456-67.: [48] Comte, C., et al., Mitochondrial targeting of recombinant RNAs modulates the level of a heteroplasmic mutation in human mitochondrial DNA associated with Kearns Sayre Syndrome. *Nucleic Acids Res*, 2013. 41(1): p. 418-33.: [49] Zelenka, J., et al., Import of desired nucleic acid sequences using addressing motif of mitochondrial ribosomal 5S-rRNA for fluorescent in vivo hybridization of mitochondrial DNA and RNA. *J Bioenerg Biomembr*, 2014. 46(2): p. 147-56.: [50] Konermann, S., et al., Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*, 2014.: [51] Fu, Y., et al., Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*, 2014. 32(3): p. 279-84.

[0272] Example 2. *Mitochondrial DNA Manipulation Using CRISPR/Cas9 Genome Editing Technology*

[0273] **Specific Aims:** The CRISPR single-guide RNA (sgRNA) and Cas9 protein are engineered for trafficking into the mitochondrial matrix for genome editing of mitochondrial DNA (mtDNA). Our overarching goal is to develop novel therapeutic strategies to treat and potentially cure mitochondrial diseases resulting from pathogenic mtDNA mutations that exist in heteroplasmy.

[0274] Aim 1: To create sgRNAs that efficiently localize into the mitochondrial matrix.

[0275] Aim 2: To evaluate and optimize mito-sgRNA specificity for mutant mtDNA target.

[0276] Aim 3: To restore oxidative phosphorylation in cybrid cells with mitoCRISPR.

[0277] **Significance:** Mitochondria are unique organelles that are the powerhouse of the cell and carry its own genomic content. Mitochondrial DNA (mtDNA) is a double-stranded circular molecule that encodes 37 genes, 24 of which are necessary for mtDNA translation (2 ribosomal RNAs, 22 transfer RNAs) and 13 subunits of the respiratory chain (complex I, III, IV and V) critical for producing energy in the form of ATP. MtDNA is present in hundreds to thousands of copies inside the cell and nucleotide polymorphisms produce a state of heteroplasmy. Random genetic drift during cell division can result in some cells or tissues containing higher loads of pathogenic mtDNA. Once a critical mutation threshold has been reached, cells exhibit a bioenergetic defect due to dysfunction

of the respiratory chain. Many mitochondrial diseases lead to devastating disorders of encephalomyopathies wherein tissues with high metabolic demands, such as musculoskeletal and neuronal tissues, are severely affected.

[0278] Strategies aimed at eliminating mutant mtDNA have shown to be effective in shifting heteroplasmy towards lower mutation load and rescuing cellular metabolic defects, thus establishing the proof-of-principle of the propose strategy(1). There are limited DNA repair mechanisms in mammalian mitochondria, and given the high redundancy of mitochondrial genome in the cell, clearance of mtDNA is a predominant mechanism in protecting the fidelity of mtDNA in mammalian cells(2, 3). As a result, targeting of restriction endonucleases or homing endonucleases, such as transcription activator-like effector nucleases (TALENs) and zinc-finger nucleases (ZFNs), to mitochondria with high levels of heteroplasmy resulted in successful depletion of mutant mtDNA and rescue of metabolic defects(2, 4-8). Furthermore, these gene editing modalities have been utilized for inducing heteroplasmic shift in the germline of murine models and human oocytes with minimal adverse effects on cell development or animal fitness suggesting that the strategy is a viable clinical therapy(9). However, the generalizability of these tools for clinical therapy is limited. Usage of restriction endonucleases requires a specific mutation that creates a compatible restriction site, and thus it is not a generalizable technique for the vast amount of characterized mtDNA mutations. Homing endonucleases such as TALENs and ZFNs require mitochondrial import of large bulky protein motifs for sequence recognition and often result in insufficient expression and poor localization in mitochondria. Although mitochondrial replacement by means of spindle-chromosomal complex transfer in oocytes(10) or somatic cell nuclear transfer in pluripotent stem cells(11) has shown promise, recent studies showed that mismatched mitochondria in nuclear-transfer-derived embryonic stem cells (NT-ESCs) can trigger an adaptive immune response and cause immune rejection of the graft(12).

[0279] **Innovation:** The system described herein adapts the CRISPR genome editing technology towards manipulating mtDNA. Our goal is to engineer this technology for targeting heteroplasmic mtDNA mutations in cybrid models of mitochondrial diseases. The development of mitoCRISPR, as we have termed it, would offer greater versatility and specificity over other mitochondrial-localizing genome editing platforms (e.g., mito-ZFN, mito-TALEN, etc.), since mitoCRISPR could be tailored to target any of the >200 documented mtDNA mutations with sequence-specific sgRNAs.

[0280] **Results:** We have successfully targeted the *S. pyogenes* Cas9 protein into the mitochondrial matrix by removing the nuclear localization signals and adding an N-terminal mitochondrial targeting sequence derived from the cytochrome *c* oxidase subunit VIII (Cox8). We have generated a lentiviral vector with mitoCas9-Flag protein driven by the SFFV promoter and a U6 RNA polymerase III promoter that drives expression of the mitochondrial-localizing sgRNA (mito-sgRNA) (FIG. 1A). Confocal analysis shows highly co-localized signal between mitoCas9-Flag protein with mitochondrial outer membrane marker Tom20 (FIG. 1C).

[0281] Mitochondrial fractionation also demonstrates an enrichment of Cas9-flag in the mitochondrial fraction (FIG. 1B). Given the new discovery that Cpf1 is another Class II (type V) CRISPR protein with similar genome editing efficiency in mammalian cells relative to Cas9(13), we are now constructing mitoCpf1 using a similar strategy.

[0282] To create sgRNAs that efficiently localize into the mitochondrial matrix: We are developing strategies for directing the mitochondrial import of the sgRNA using rationale design. Mechanisms for RNA import in mammalian mitochondria have not been defined, but several small RNAs have been isolated from the mitochondrial transcriptome, including the 5S ribosomal RNA, RNaseP and MRP RNA components (reviewed in (14)). Importantly, the delivery of exogenous mRNA or tRNA into mitochondria has been demonstrated by attaching these hairpin loops to the exogenous RNA (15-17). Additionally, a modified γ domain of 5S rRNA exhibits enhanced import of RNA into the mitochondria (18, 19). We will verify the efficacy of these mitochondrial localization loops (mitoloop) for mitochondrial import of the sgRNA (FIG. 2A). We have designed insertion sites of various mitoloops at positions in the sgRNA that have minimal interactions with Cas9, as seen in Fig. 2B. In particular, we have selected the tetraloop and the stem loop 2 to be replaced by mitoloops as mutations in these regions bear no impact on Cas9 binding or activity(20). The import of sgRNA into the mitochondria will be verified by imaging analysis and quantitative reverse transcription PCR (qRT-PCR). Fluorescently labeled sgRNA with various mitoloops will be introduced to mammalian cells via transfection or electroporation techniques. Co-localization analysis will be performed between the sgRNA signal and a mitochondrial marker as a measure of sgRNA import efficiency. Furthermore, we will quantify sgRNA content in purified mitoplast, formed by selective dissolution of mitochondrial outer membrane, by qRT-PCR.

[0283] To evaluate and optimize mito-sgRNA specificity for mutant mtDNA target: There are over 200 mutations in mtDNA associated with devastating encephalomyopathies, and most of these mutations are single point mutations. We are optimizing the specificity of CRISPR genome editing by examining the minimum length of sgRNA to maximize on-target DNA recognition. Truncation of the complementarity region of the sgRNA to 17-18 nt can significantly reduce off target cleavage with minimal effects on on-target recognition (25). To examine the specificity of CRISPR towards mtDNA targets, we will develop a nuclear assay with dual chemiluminescence or fluorescence to assess cleavage activity at the wildtype or mutant DNA target in the nucleus with nuclear Cas9. Small regions of the mtDNA with wildtype sequence or a single point mutation will be fused in frame with a chemiluminescence reporter. Cleavage of the target should result in insertions/deletions (indels) that may disrupt the downstream translation of the reporter and result in a loss of signal. Various sgRNA lengths will be tested for DNA recognition specificity. Furthermore, the protospacer adjacent motif (PAM) domain is critical region in Cas9-mediated DNA melting and RNA-DNA heteroduplex formation. Several mutation variants of Cas9 have been generated with expanded repertoire of PAM recognition sites. We will determine if utilizing mutations that create a new PAM domain will enhance the specificity of CRISPR technology for targeting single point mutations in mtDNA.

[0284] To restore oxidative phosphorylation in cybrid cells with mitoCRISPR: After we have determined the mitoloop design with highest import efficiency, we will apply the mitoCRISPR technology with mito-sgRNA in cellular models of mitochondrial diseases. Cytoplasmic hybrids (cybrids) have become an invaluable resource for studying the mechanisms and testing therapies for mitochondrial diseases due to the inability to transform mtDNA. Cybrids are derived from enucleated fibroblasts from patients that carry mitochondria with heteroplasmic or homoplasmic mutations. These cells are subsequently fused with an immortalized osteosarcoma cell line that has been depleted of mitochondria. We have obtained homoplasmic cybrid cell lines that carry individual mtDNA mutations in complex I, III, IV, and IV and a heteroplasmic cybrid for the MERRF mutation to test gene editing by mitoCRISPR. Cybrids will be transduced by lentiviral vector carrying the mito-sgRNA with the appropriate mitoloop and mitoCas9. Depletion of mutant mtDNA will be assessed by quantitative PCR and possibly by digital droplet PCR (ddPCR) for

greater sensitivity. Sequencing of mtDNA will be required to verify the efficiency of target cleavage. Mitochondrial respiration will be measured to assess functional rescue of decreasing mutation load.

[0285] In order to engineer the CRISPR/Cas9 system to operate on the mtDNA genome, targeting of the protein component (Cas9) and the RNA component (sgRNA) into the mitochondrial matrix is desirable. An exemplary construct to achieve this end is set forth in FIG. 2.

[0286] References (Example 2)

[0287] [1] Lightowlers RN, Taylor RW, Turnbull DM. Mutations causing mitochondrial disease: What is new and what challenges remain? *Science*. 2015;349(6255):1494-9. Epub 2015/09/26. doi: 10.1126/science.aac7516. PubMed PMID: 26404827.; [2] Alexeyev M, Shokolenko I, Wilson G, LeDoux S. The maintenance of mitochondrial DNA integrity--critical analysis and update. *Cold Spring Harb Perspect Biol*. 2013;5(5):a012641. doi: 10.1101/cshperspect.a012641. PubMed PMID: 23637283; PubMed Central PMCID: PMC3632056.; [3] Kazak L, Reyes A, Holt IJ. Minimizing the damage: repair pathways keep mitochondrial DNA intact. *Nat Rev Mol Cell Biol*. 2012;13(10):659-71. doi: 10.1038/nrm3439. PubMed PMID: 22992591.; [4] Bacman SR, Williams SL, Pinto M, Peralta S, Moraes CT. Specific elimination of mutant mitochondrial genomes in patient-derived cells by mitoTALENs. *Nat Med*. 2013;19(9):1111-3. Epub 2013/08/06. doi: 10.1038/nm.3261. PubMed PMID: 23913125; PubMed Central PMCID: PMC4153471.; [5] Bayona-Bafaluy MP, Blits B, Battersby BJ, Shoubridge EA, Moraes CT. Rapid directional shift of mitochondrial DNA heteroplasmy in animal tissues by a mitochondrially targeted restriction endonuclease. *Proc Natl Acad Sci U S A*. 2005;102(40):14392-7. doi: 10.1073/pnas.0502896102. PubMed PMID: 16179392; PubMed Central PMCID: PMC1242285.; [6] Gammage PA, Rorbach J, Vincent AI, Rebar EJ, Minczuk M. Mitochondrially targeted ZFNs for selective degradation of pathogenic mitochondrial genomes bearing large-scale deletions or point mutations. *EMBO Mol Med*. 2014;6(4):458-66. Epub 2014/02/26. doi: 10.1002/emmm.201303672. PubMed PMID: 24567072; PubMed Central PMCID: PMC3992073.; [7] Hashimoto M, Bacman SR, Peralta S, Falk MJ, Chomyn A, Chan DC, et al. MitoTALEN: A General Approach to Reduce Mutant mtDNA Loads and Restore Oxidative Phosphorylation Function in Mitochondrial Diseases. *Mol Ther*. 2015;23(10):1592-9. doi: 10.1038/mt.2015.126. PubMed PMID: 26159306.; [8] Minczuk M, Kolasinska-Zwierz P, Murphy MP, Papworth MA. Construction and testing of engineered zinc-finger proteins for sequence-specific modification of mtDNA. *Nat Protoc*. 2010;5(2):342-56. Epub

- 2010/02/06. doi: 10.1038/nprot.2009.245. PubMed PMID: 20134433.; [9] Reddy P, Ocampo A, Suzuki K, Luo J, Bacman SR, Williams SL, et al. Selective elimination of mitochondrial mutations in the germline by genome editing. *Cell*. 2015;161(3):459-69. Epub 2015/04/25. doi: 10.1016/j.cell.2015.03.051. PubMed PMID: 25910206.; [10] Tachibana M, Sparman M, Sritanaudomchai H, Ma H, Clepper L, Woodward J, et al. Mitochondrial gene replacement in primate offspring and embryonic stem cells. *Nature*. 2009;461(7262):367-72. Epub 2009/08/28. doi: 10.1038/nature08368. PubMed PMID: 19710649; PubMed Central PMCID: PMC2774772.; [11] Ma H, Folmes CD, Wu J, Morey R, Mora-Castilla S, Ocampo A, et al. Metabolic rescue in pluripotent cells from patients with mtDNA disease. *Nature*. 2015;524(7564):234-8. Epub 2015/07/16. doi: 10.1038/nature14546. PubMed PMID: 26176921.; [12] Deuse T, Wang D, Stubbendorff M, Itagaki R, Grabosch A, Greaves LC, et al. SCNT-derived ESCs with mismatched mitochondria trigger an immune response in allogeneic hosts. *Cell stem cell*. 2015;16(1):33-8. Epub 2014/12/04. doi: 10.1016/j.stem.2014.11.003. PubMed PMID: 25465116.; [13] Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, Essletzbichler P, et al. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*. 2015;163(3):759-71. Epub 2015/10/01. doi: 10.1016/j.cell.2015.09.038. PubMed PMID: 26422227; PubMed Central PMCID: PMC4638220.; [14] Schneider A. Mitochondrial tRNA import and its consequences for mitochondrial translation. *Annu Rev Biochem*. 2011;80:1033-53. Epub 2011/03/23. doi: 10.1146/annurev-biochem-060109-092838. PubMed PMID: 21417719.; [15] Wang G, Chen HW, Oktay Y, Zhang J, Allen EL, Smith GM, et al. PNPASE regulates RNA import into mitochondria. *Cell*. 2010;142(3):456-67. Epub 2010/08/10. doi: S0092-8674(10)00725-7 10.1016/j.cell.2010.06.035. PubMed PMID: 20691904; PubMed Central PMCID: PMC2921675.; [16] Comte C, Tonin Y, Heckel-Mager AM, Boucheham A, Smirnov A, Aure K, et al. Mitochondrial targeting of recombinant RNAs modulates the level of a heteroplasmic mutation in human mitochondrial DNA associated with Kearns Sayre Syndrome. *Nucleic Acids Res*. 2013;41(1):418-33. doi: 10.1093/nar/gks965. PubMed PMID: 23087375; PubMed Central PMCID: PMC3592399.; [17] Tonin Y, Heckel AM, Vysokikh M, Dovydenko I, Meschaninova M, Rotig A, et al. Modeling of antigenomic therapy of mitochondrial diseases by mitochondrially addressed RNA targeting a pathogenic point mutation in mitochondrial DNA. *J Biol Chem*. 2014;289(19):13323-34. doi: 10.1074/jbc.M113.528968. PubMed PMID: 24692550; PubMed Central PMCID: PMC4036341.; [18] Smirnov A, Tarassov I, Mager-Heckel AM, Letzelter M,

Martin RP, Krashennnikov IA, et al. Two distinct structural elements of 5S rRNA are needed for its import into human mitochondria. *RNA*. 2008;14(4):749-59. doi: 10.1261/rna.952208. PubMed PMID: 18314502; PubMed Central PMCID: PMC2271358.; [19] Zelenka J, Alan L, Jaburek M, Jezek P. Import of desired nucleic acid sequences using addressing motif of mitochondrial ribosomal 5S-rRNA for fluorescent in vivo hybridization of mitochondrial DNA and RNA. *J Bioenerg Biomembr*. 2014;46(2):147-56. doi: 10.1007/s10863-014-9543-2. PubMed PMID: 24562889.; [20] Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature*. 2014. Epub 2014/12/11. doi: 10.1038/nature14136. PubMed PMID: 25494202.; [21] Zhou J, Satheesan S, Li H, Weinberg MS, Morris KV, Burnett JC, et al. Cell-specific RNA aptamer against human CCR5 specifically targets HIV-1 susceptible cells and inhibits HIV-1 infectivity. *Chem Biol*. 2015;22(3):379-90. Epub 2015/03/11. doi: 10.1016/j.chembiol.2015.01.005. PubMed PMID: 25754473; PubMed Central PMCID: PMC4369413.; [22] Zhou J, Tiemann K, Chomchan P, Alluin J, Swiderski P, Burnett J, et al. Dual functional BAFF receptor aptamers inhibit ligand-induced proliferation and deliver siRNAs to NHL cells. *Nucleic Acids Res*. 2013;41(7):4266-83. Epub 2013/03/09. doi: 10.1093/nar/gkt125. PubMed PMID: 23470998; PubMed Central PMCID: PMC3627597.; [23] Takahashi M, Burnett JC, Rossi JJ. Aptamer-siRNA chimeras for HIV. *Advances in experimental medicine and biology*. 2015;848:211-34. Epub 2015/03/12. doi: 10.1007/978-1-4939-2432-5_11. PubMed PMID: 25757623.; [24] Burnett JC, Rossi JJ. RNA-based therapeutics: current progress and future prospects. *Chem Biol*. 2012;19(1):60-71. Epub 2012/01/31. doi: S1074-5521(11)00459-5 [pii] 10.1016/j.chembiol.2011.12.008. PubMed PMID: 22284355; PubMed Central PMCID: PMC3269031.; [25] Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat Biotechnol*. 2014;32(3):279-84. doi: 10.1038/nbt.2808. PubMed PMID: 24463574; PubMed Central PMCID: PMC3988262.

[0288] *Example 3. Exemplary delivery vehicle.*

[0289] We are currently testing electroporation of the Cas9 mRNA and sgRNA (Amaxa Electroporator for laboratory scale and MaxCyte for clinical/GMP scale); liposomal formulation of Cas9 mRNA and sgRNA (encoded from plasmid DNA, or as separate RNA molecules, or as plasmid DNA for mtCas9 and in vitro transcribed RNA for the sgRNA); and lentiviral delivery in a vector

that expresses both the mtCas9 gene and the sgRNA. We have successfully used liposomes (Lipofectamine 2000) to deliver a plasmid that encodes mtCas9 together with in vitro transcribed sgRNA (FIG. 4). Nanoparticles or dendrimers are used for delivery of Cas9 or sgRNA.

[0290] *Example 4. Methods.*

5 [0291] **Cybrid cell culturing**

[0292] Cybrids were maintained in DMEM media (ThermoFisher #12800082) containing 10% fetal bovine serum, 100 U/mL penicillin/streptomycin, high glucose (25 mM), 1 mM pyruvate, 4 mM L-Glutamine and uridine 50 µg/ml. The cytochrome b mutants carry with homoplasmic deletion of 4 nucleotides at position 14787. They were generated from 143B/206 osteosarcoma
10 nuclear background and described previously (2). Cells were cultured in a 37°C incubator with CO₂ maintained at 5%.

[0293] **Stable cell lines**

[0294] Lentiviral vectors were packaged in 293T cells and purified by ultracentrifugation as described previously (3). Hela and cybrids cells were transduced with bicistronic lentiviral vectors
15 that expressed both mitoCRISPR and TagRFP. Spin transduction was performed at 300 g for 30 min in the presence of 4 µg/uL polybrene. Cells were transduced at a multiplicity of infection less than 0.5. Viral containing media were removed 6 hours post-transduction. Cells were subsequently sorted for TagRFP by flow cytometry activated cell sorting (FACS).

[0295] **Mitoplast isolation and Taqman Quantitative Reverse Transcription-PCR**

20 [0296] The enrichment of sgRNA in mitochondria was assessed by quantifying the amount of sgRNA in purified mitoplasts relative to total cytoplasm. Mitochondria were isolated from cells by differential centrifugation and treated with digitonin 1mg/mL on ice for 10 min with intermittent agitation to generate mitoplasts. Cytosolic RNA contamination was removed by treatment with RnaseA 100 µg/mL and micrococcal S7 nuclease 300U (ThermoFisher) for 30 min at room
25 temperature. Mitoplast and cytoplasmic RNA were extracted by Trizol. Primers and TaqMan probes used for sgRNA, mitochondrial housekeeping RNA cytochrome B (cytB), and cytoplasm housekeeping RNA beta actin (β -actin) are listed below. Data were first analyzed by ΔΔCt method relative to the “no mitoloop control” as well as either the cytB or bactin control RNAs. Data were

then normalized to combine $\Delta\Delta C_t$ for both housekeeping genes into a single numerical value by computing the Euclidean distance from the “no mitoloop control” (4, 5). Hence, the value of the “no mitoloop control” was set to 1.0.

Table 9	
Primers	Taqman Probe
Cyt B Fwd: GCCTATATTACGGATCATTTCTCT ACT (SEQ ID NO: 113) Cyt B Rev: GCCTATGAAGGCTGTTGCTATAGT (SEQ ID NO: 114)	CytB probe: CCTGAAACATCGGCATTATCCTCCTGCT (SEQ ID NO: 117)
β -actin Fwd: ACCTGACTGACTACCTCATGAAG ATCCTCACCGA (SEQ ID NO: 115) β -actin Rev: GGAGCTGGAAGCAGCCGTGGCCA TCTCTTGCTCGAA (SEQ ID NO: 116)	β -actin: AGCGGGAAATCGTGCGTGACATTA (SEQ ID NO: 118)

5

[0297] FIGS. 6A-6C show that CRISPR endonuclease AsCpf1 can be engineered to localize to the mitochondria, which is termed mitoAsCpf1. FIGS. 1A-1C and FIGS. 4A-4C had demonstrated this for an engineered version of spCas9 (mitoCas9), which is the most commonly used CRISPR endonuclease. Thus, the mitoCRISPR technology extends beyond Cas9-based CRISPR systems, and applies to other CRISPR systems.

10

[0298] FIG. 7 shows a functional assay for mitoCRISPR activity with mitoCas9 and engineered sgRNAs with different mitoloops. The functional readout is the mitochondrial morphology, which is visibly perturbed in the sgRNAs with mitoloops.

15

[0299] FIG. 8 shows the enrichment of sgRNAs into mitochondria. This experiment confirms that sgRNAs with mitoloops (in particular the LDF mitoloop) can be trafficked to mitochondria with increased efficiency. This is a novel demonstration. This data confirmed achieving mitoCRISPR using the sgRNA-mitoloop design.

[0300] References (Example 4)

- [0301]** [1] Kolesnikova O, et al. (2010) Selection of RNA aptamers imported into yeast and human mitochondria. RNA 16(5):926-941.; [2] Rana M, de Coo I, Diaz F, Smeets H, & Moraes CT (2000) An out-of-frame cytochrome b gene deletion from a patient with parkinsonism is associated with
5 impaired complex III assembly and an increase in free radical production. Annals of neurology 48(5):774-781.; [3] Burnett JC, Miller-Jensen K, Shah PS, Arkin AP, & Schaffer DV (2009) Control of stochastic gene expression by host factors at the HIV promoter. PLoS pathogens 5(1):e1000260.; [4] Bergkvist A, et al. (2010) Gene expression profiling--Clusters of possibilities. Methods 50(4):323-335.; [5] Wolf AR & Mootha VK (2014) Functional genomic analysis of human
10 mitochondrial RNA processing. Cell Rep 7(3):918-931.

WHAT IS CLAIMED IS:

- 1 1. A composition comprising a delivery vehicle and a protein comprising a
2 mitochondrial localization amino acid sequence covalently attached to an RNA-guided DNA
3 endonuclease enzyme, wherein said protein is bound to said delivery vehicle.
- 1 2. The composition of claim 1, wherein said delivery vehicle is a nanoparticle or
2 a lipid particle or a viral vector.
- 1 3. The composition of claim 1, wherein said protein is encapsulated within said
2 delivery vehicle.
- 1 4. The composition of claim 1, wherein said mitochondrial localization amino
2 acid sequence is N-terminal to said RNA-guided DNA endonuclease enzyme.
- 1 5. The composition of claim 1, wherein said mitochondrial localization amino
2 acid sequence is a cytochrome c oxidase subunit VIII (Cox8) sequence.
- 1 6. The composition of claim 1, wherein said RNA-guided DNA endonuclease
2 enzyme is Cas9 or Cpf1 or a Class II CRISPR endonuclease.
- 1 7. The composition of claim 1, wherein said RNA-guided DNA endonuclease
2 enzyme has no nuclear localization sequence.
- 1 8. The composition of claim 4, wherein said Cas9 is a Cas9 variant, wherein said
2 Cas9 variant has one or more mutations that increase its binding specificity to PAM compared to
3 wild type Cas9.
- 1 9. A composition comprising a delivery vehicle and a nucleic acid encoding said
2 protein of one of claims 1 to 8.
- 1 10. A protein comprising a mitochondrial localization amino acid sequence
2 covalently attached to Cpf1 or a Class II CRISPR endonuclease or a Cas9 variant.
- 1 11. A nucleic acid encoding said protein of claim 10.

1 12. A nucleic acid comprising a mitochondrial import sequence and a single-guide
2 RNA (sgRNA) sequence.

1 13. The nucleic acid of claim 12, wherein said mitochondrial import sequence
2 comprises a D loop, an F loop, an MRP loop, an RNP loop, a γ 5s loop, or any combination thereof.

1 14. The nucleic acid of claim 12, wherein said sgRNA sequence comprises a
2 nucleic acid sequence that is complementary to a mitochondrial DNA (mtDNA) target sequence.

1 15. The nucleic acid of claim 14, wherein said mtDNA target sequence comprises
2 at least one mutation or deletion.

1 16. The nucleic acid of claim 15, wherein said mtDNA target sequence is 16-24
2 nucleotides in length.

1 17. The nucleic acid of claim 12, wherein said sgRNA sequence comprises a
2 transactivating crRNA (tracrRNA) sequence.

1 18. The nucleic acid of claim 12, wherein said sgRNA sequence comprises a
2 crRNA sequence.

1 19. The nucleic acid of claim 12, wherein said sgRNA sequence comprises a
2 tracrRNA sequence and a crRNA sequence.

1 20. The nucleic acid of claim 12, wherein said sgRNA sequence comprises a
2 nuclear-encoded and mitochondrial-localizing tRNA sequence.

1 21. The nucleic acid of claim 12, wherein said nucleic acid is bound to a delivery
2 vehicle.

1 22. The nucleic acid of claim 21, wherein said delivery vehicle is a nanoparticle
2 or a lipid particle.

1 23. The nucleic acid of claim 12, wherein said nucleic acid is DNA.

- 1 24. The nucleic acid of claim 12, wherein said nucleic acid is RNA.
- 1 25. The nucleic acid of claim 24, wherein said RNA comprises one or more
2 modified nucleotides or nucleotide analogues.
- 1 26. A vector comprising the nucleic acid of claim 11, the nucleic acid of one of
2 claims 12 to 25, or a combination thereof.
- 1 27. The vector of claim 26, wherein said vector is a replication incompetent viral
2 vector.
- 1 28. The vector of claim 27, wherein said replication incompetent viral vector is a
2 replication incompetent lentiviral, adeno-associated viral, or adenoviral vector.
- 1 29. A pharmaceutical composition comprising the composition of one of claims 1
2 to 9, the protein of claim 10, the nucleic acid of one of claims 11 to 25, a vector of one of claims 26
3 to 28, and a pharmaceutically acceptable excipient.
- 1 30. A method of altering expression of at least one mitochondrial nucleic acid
2 sequence, the method comprising introducing into an eukaryotic cell the nucleic acid of one of
3 claims 12 to 25.
- 1 31. The method of claim 30, wherein said method further comprising introducing
2 into said eukaryotic cell the composition of one of claims 1 to 9, the protein of claim 10, or the
3 nucleic acid of claim 11.
- 1 32. The method of claim 30, wherein said eukaryotic cell is an oocyte.
- 1 33. The method of claim 30, wherein said eukaryotic cell is part of a fertilized
2 embryo.
- 1 34. The method of claim 30, wherein said mitochondrial nucleic acid sequence
2 comprises at least one mutation or deletion.

1 35. A method of treating a mitochondrial disorder in a subject in need thereof, the
2 method comprising administering to said subject an effective amount of the composition of one of
3 claims 1 to 9 and an effective amount of the nucleic acid of one of claims 12 to 25.

1 36. A method of treating a mitochondrial disorder in a subject in need thereof, the
2 method comprising administering to said subject an effective amount of the protein of claim 10 or
3 the nucleic acid of claim 11 and an effective amount of the nucleic acid of one of claims 12 to 25.

1 37. The method of claim 35 or claim 36, wherein said mitochondrial disorder is
2 selected from the group consisting of Myoclonic Epilepsy with Ragged Red Fibers (MERRF);
3 Mitochondrial Myopathy, Encephalopathy, Lactacidosis, and Stroke (MELAS); Maternally Inherited
4 Diabetes and Deafness (MIDD); Leber's Hereditary Optic Neuropathy (LHON); chronic progressive
5 external ophthalmoplegia (CPEO); Leigh Disease; Kearns-Sayre Syndrome (KSS); Friedreich's
6 Ataxia (FRDA); Co-Enzyme Q10 (CoQ10) Deficiency; Complex I Deficiency; Complex II
7 Deficiency; Complex III Deficiency; Complex IV Deficiency; Complex V Deficiency; other
8 myopathies; cardiomyopathy; encephalomyopathy; renal tubular acidosis; neurodegenerative
9 diseases; Parkinson's disease; Alzheimer's disease; amyotrophic lateral sclerosis (ALS); motor
10 neuron diseases; hearing and balance impairments; or other neurological disorders; epilepsy; genetic
11 diseases; Huntington's Disease; mood disorders; nucleoside reverse transcriptase inhibitors (NRTI)
12 treatment; HIV-associated neuropathy; schizophrenia; bipolar disorder; age-associated diseases;
13 cerebral vascular diseases; macular degeneration; diabetes; and cancer.

1 38. A kit, comprising a nucleic acid of one of claims 12 to 25, and a protein
2 comprising a mitochondrial localization amino acid sequence covalently attached to an RNA-guided
3 DNA endonuclease enzyme or a nucleic acid sequence encoding said protein.

1 39. The kit of claim 38, wherein said mitochondrial localization amino acid
2 sequence is N-terminal to said RNA-guided DNA endonuclease enzyme.

1 40. The kit of claim 39, wherein said mitochondrial localization amino acid
2 sequence is a cytochrome c oxidase subunit VIII (Cox8) sequence.

1 41. The kit of claim 38, wherein said RNA-guided DNA endonuclease enzyme is
2 Cas9 or Cpf1 or a Class II CRISPR endonuclease.

1 42. The kit of claim 38, wherein said RNA-guided DNA endonuclease enzyme
2 has no nuclear localization sequence.

1 43. The kit of claim 41, wherein said Cas9 is a mutant Cas9, wherein said mutant
2 Cas9 has one or more mutations that increase its binding specificity to PAM compared to wild type
3 Cas9.

1 44. The kit of claim 38, wherein said nucleic acid sequence forms part of a viral
2 vector nucleic acid sequence.

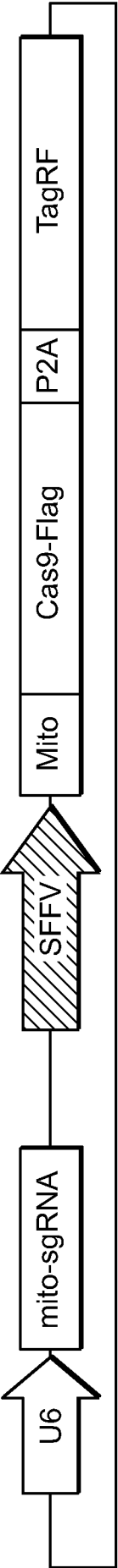


FIG. 1A

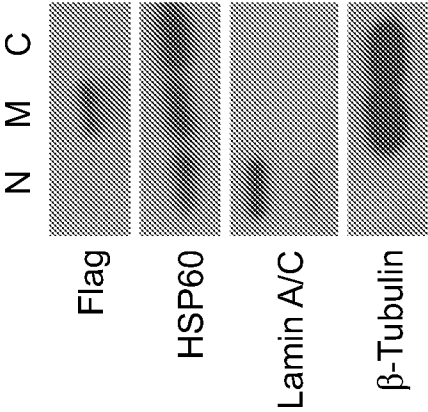


FIG. 1B

2/15

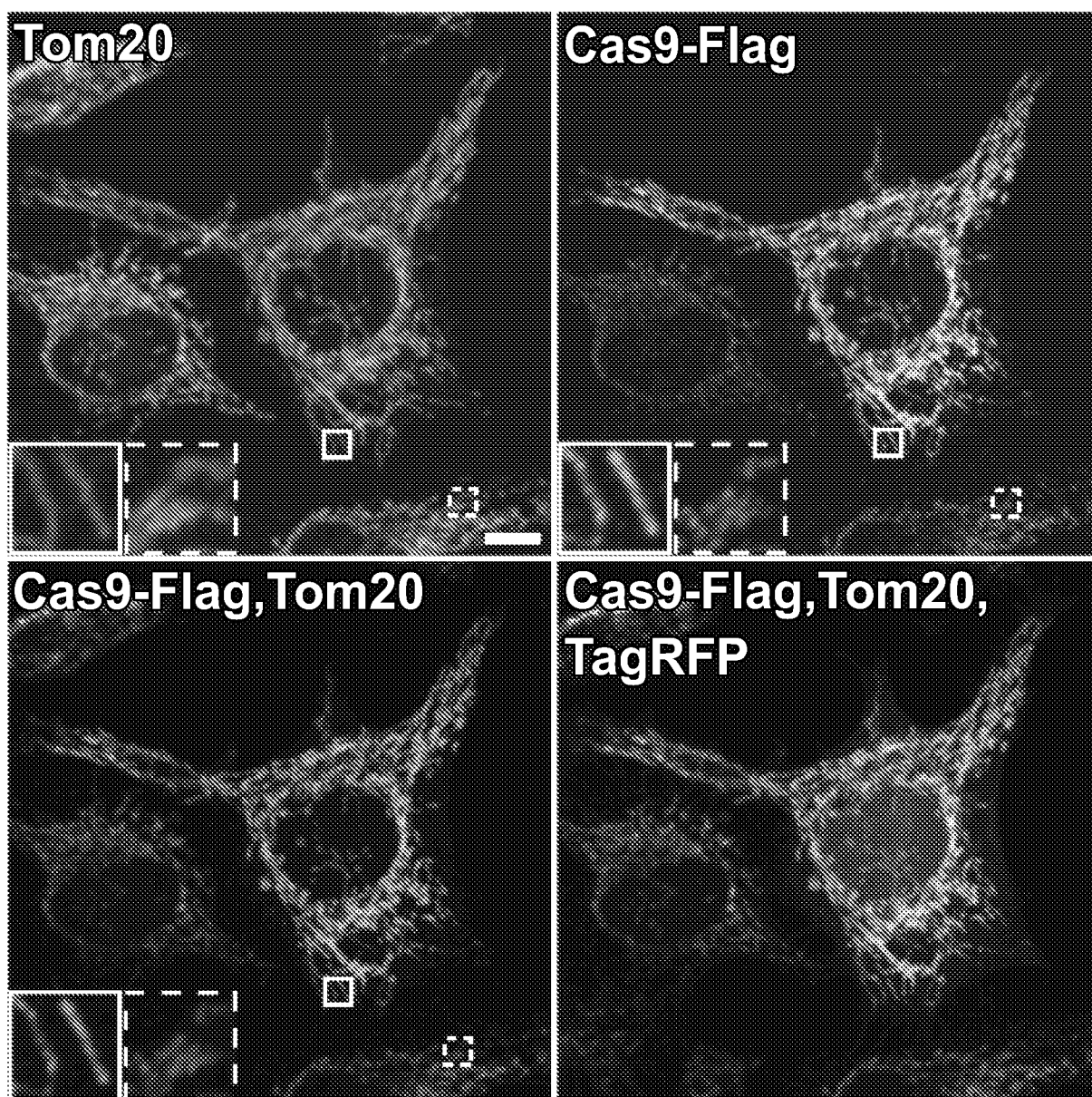
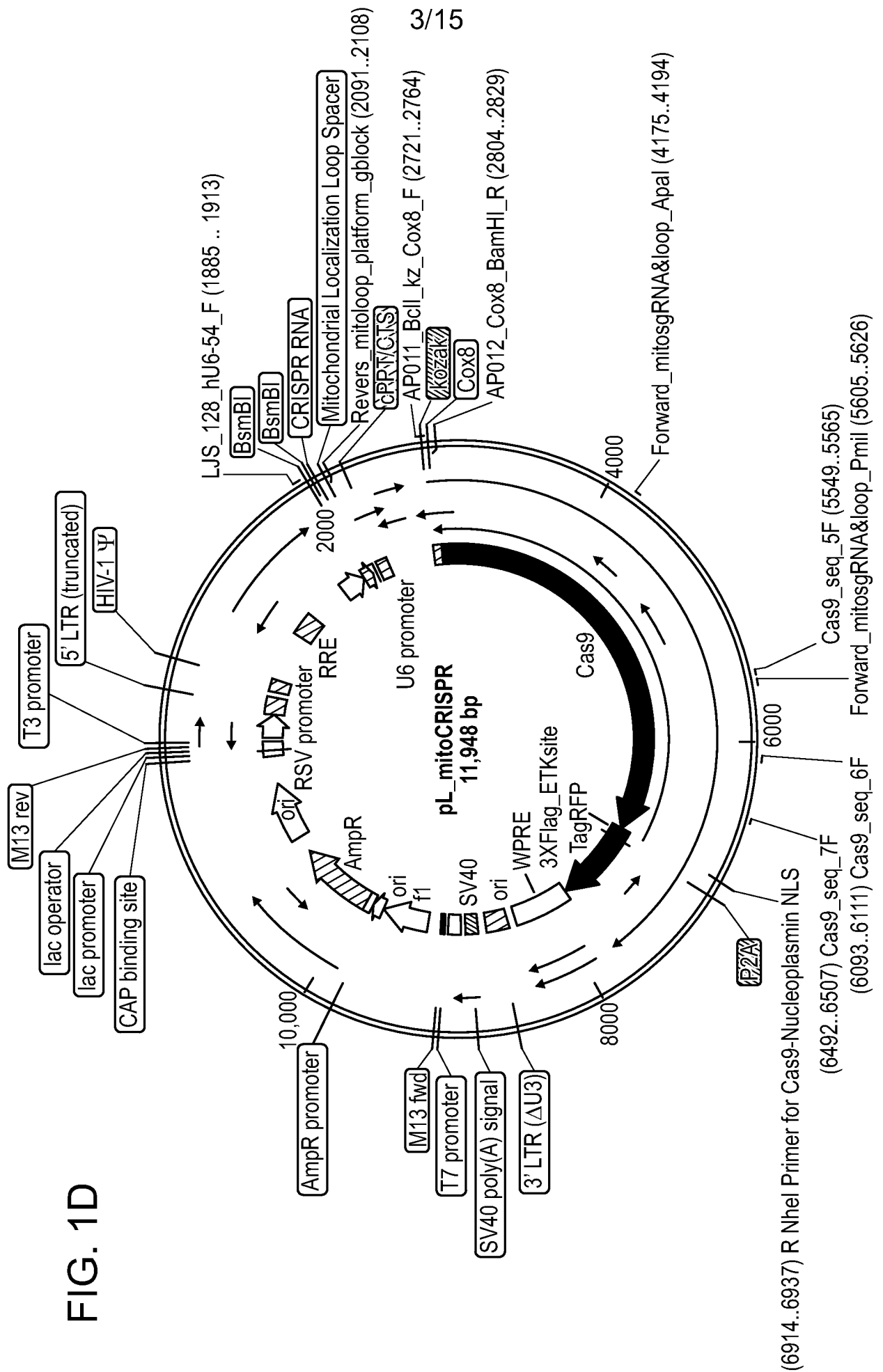
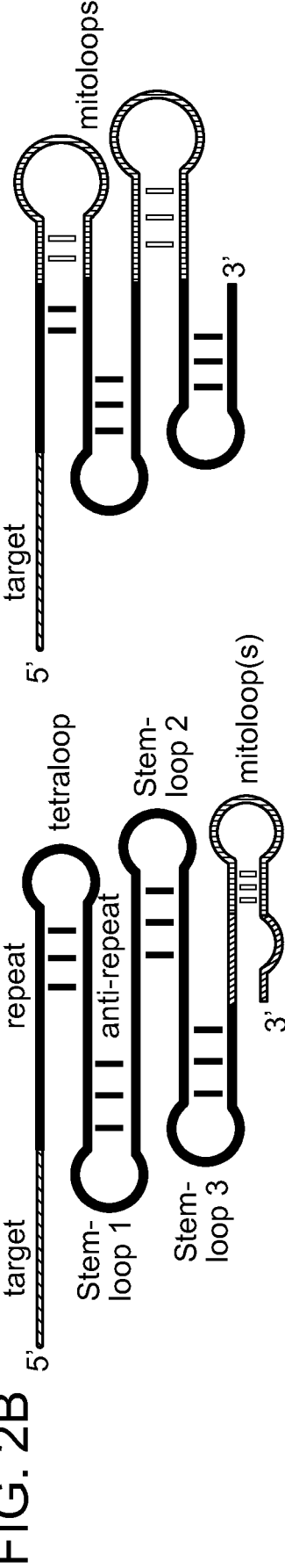
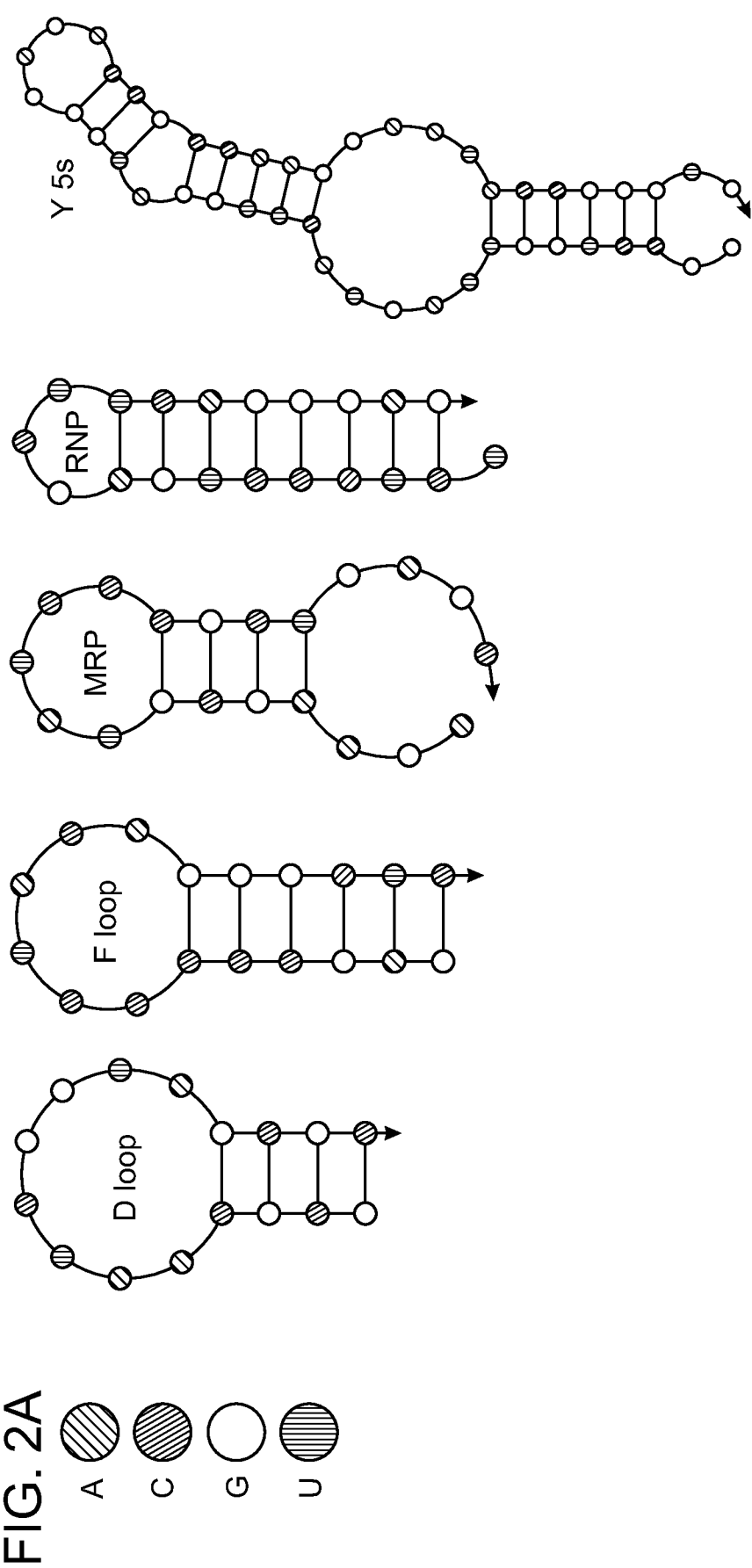


FIG. 1C

FIG. 1D





5/15

FIG. 3A

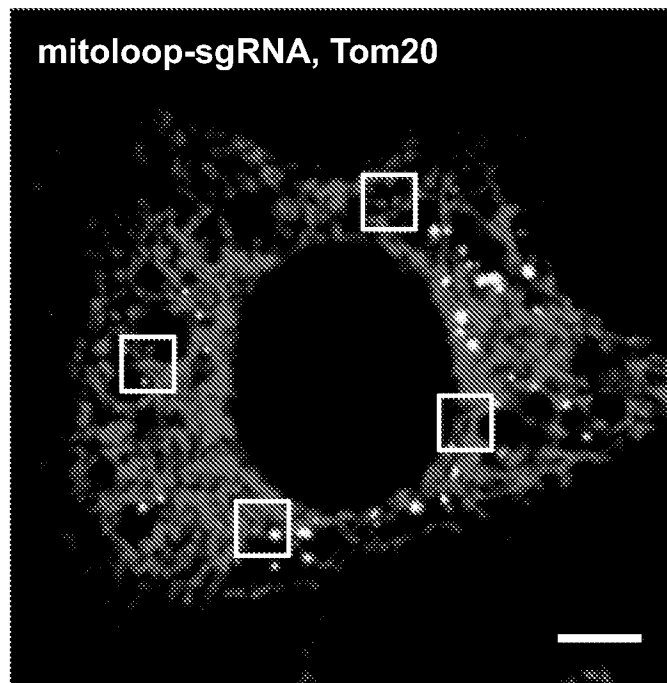
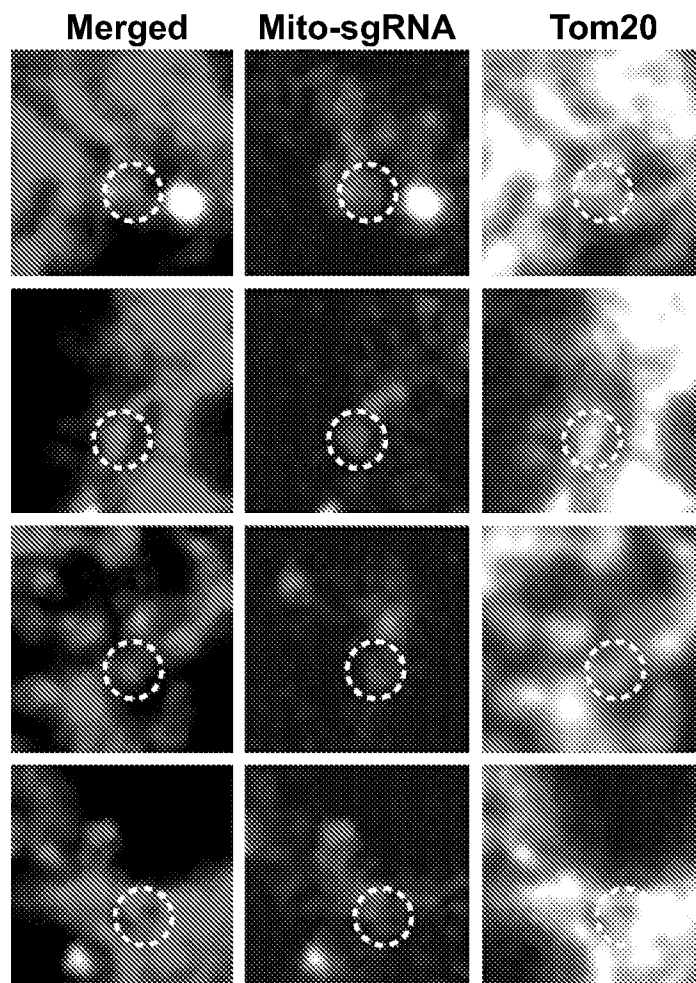


FIG. 3B



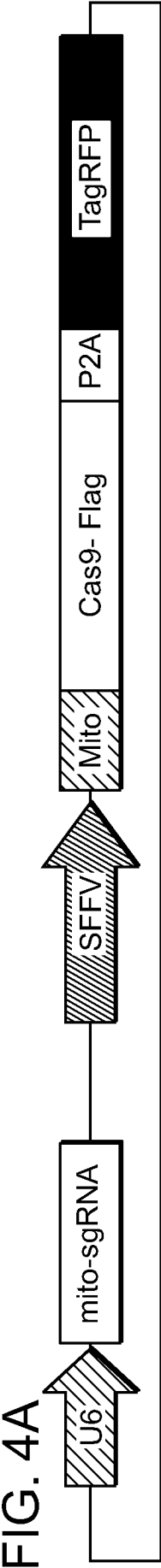


FIG. 4B

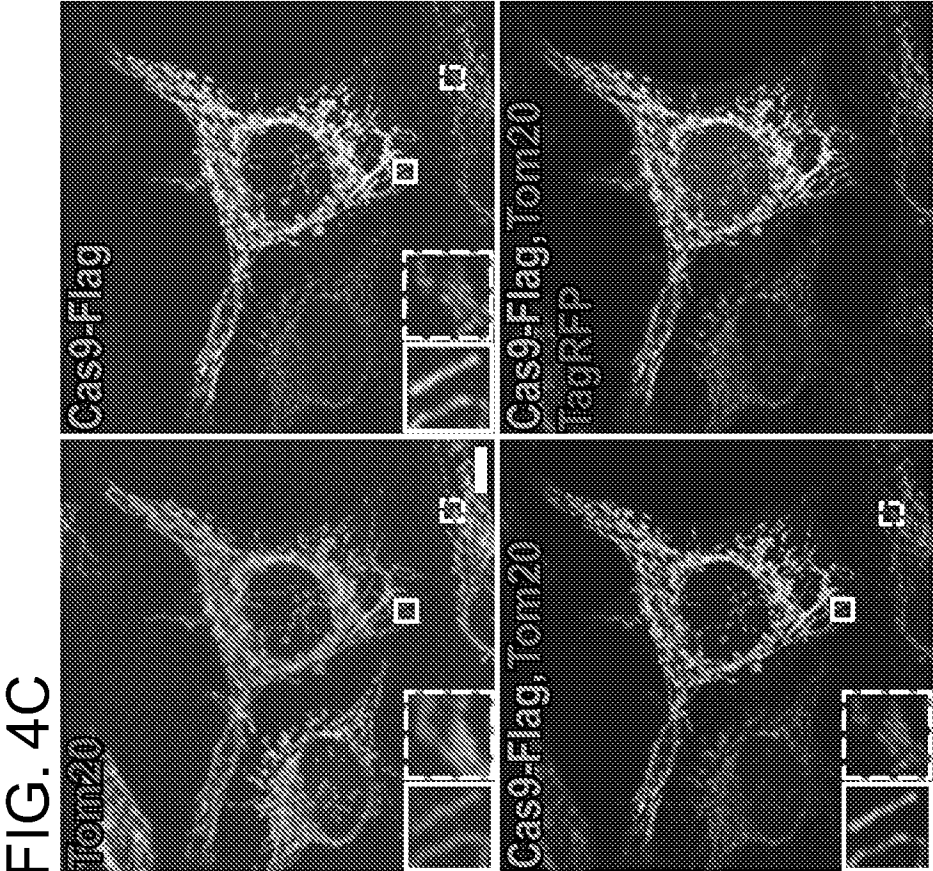
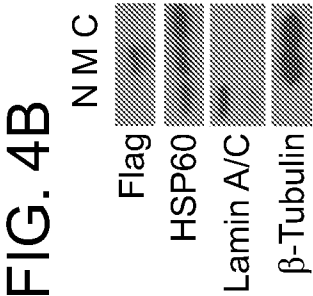


FIG. 5A

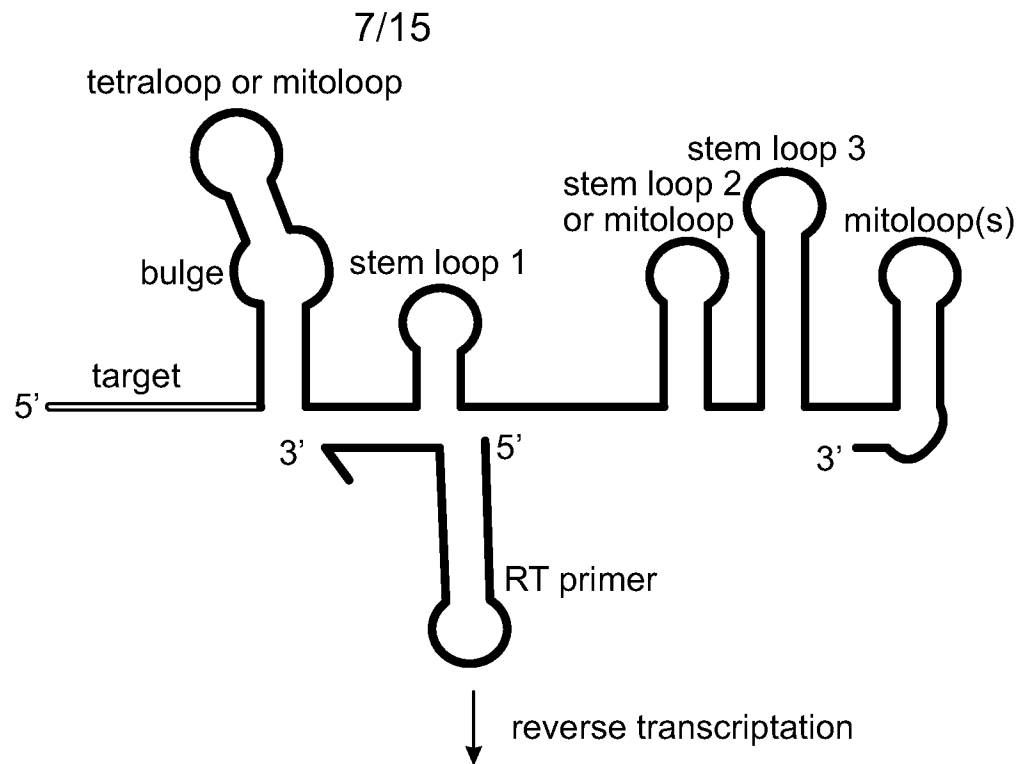


FIG. 5B

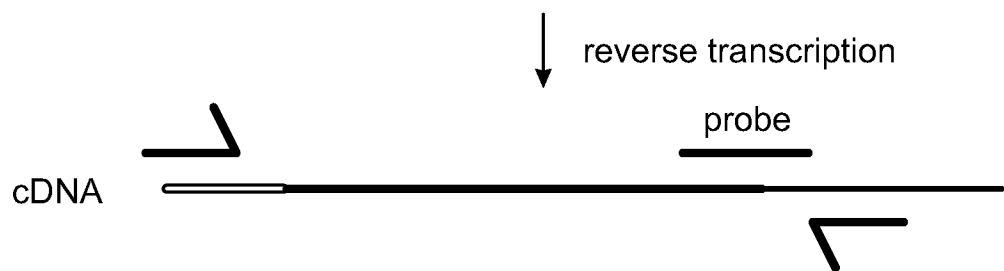
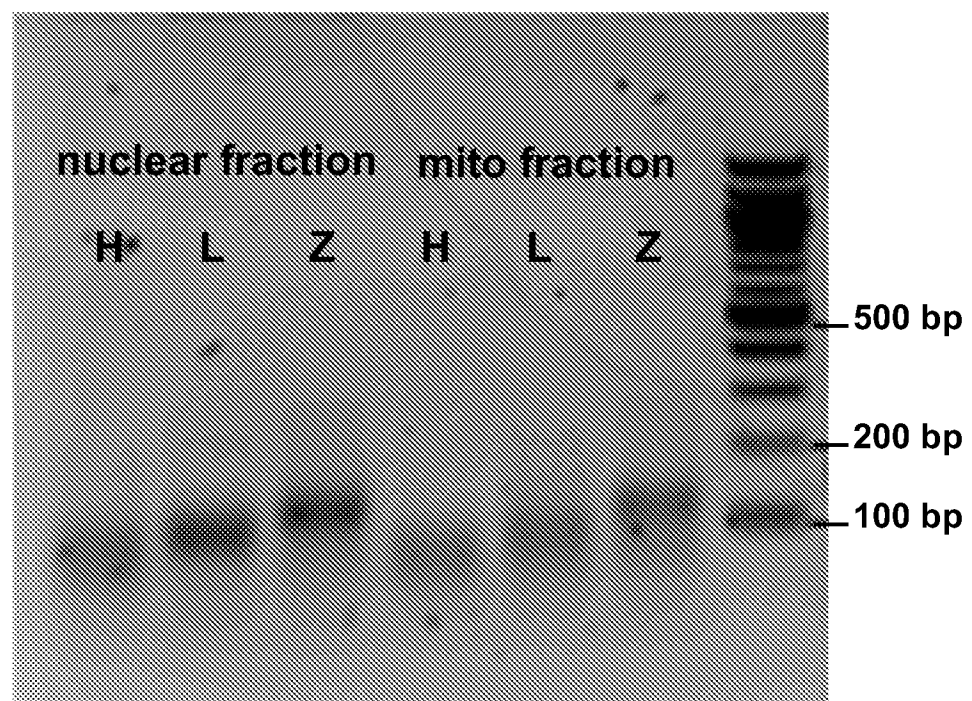


FIG. 5C



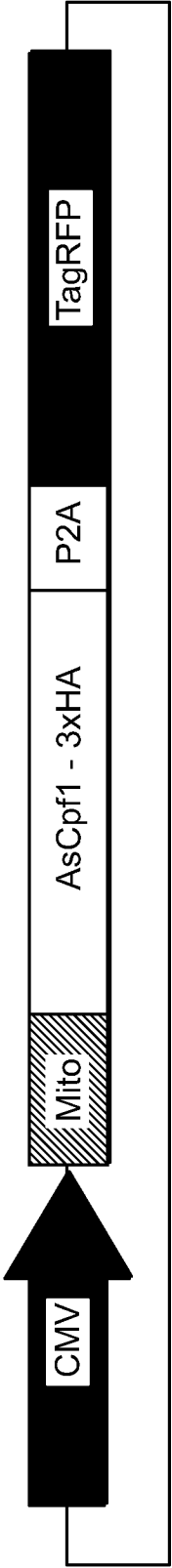


FIG. 6A

9/15

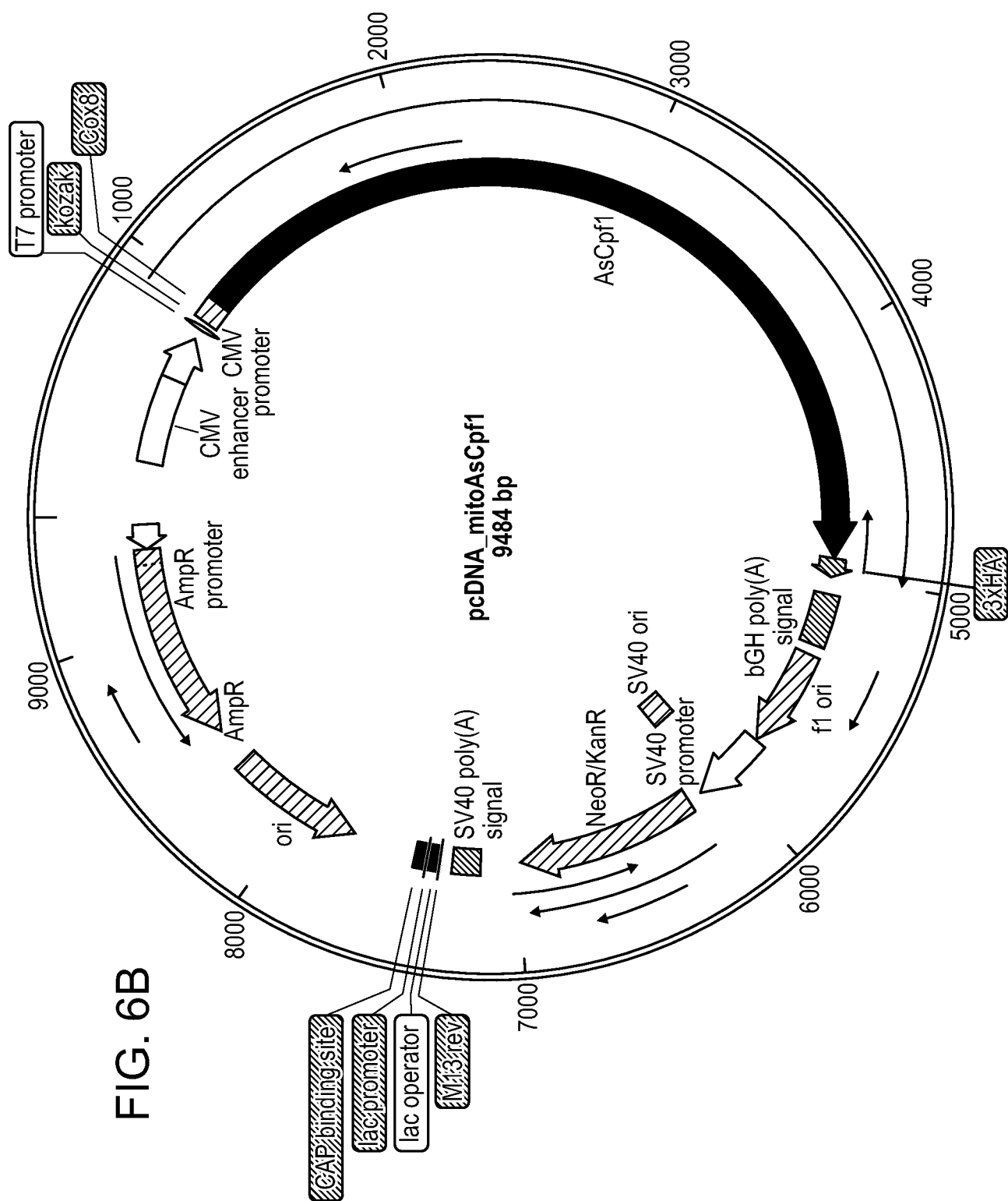


FIG. 6B

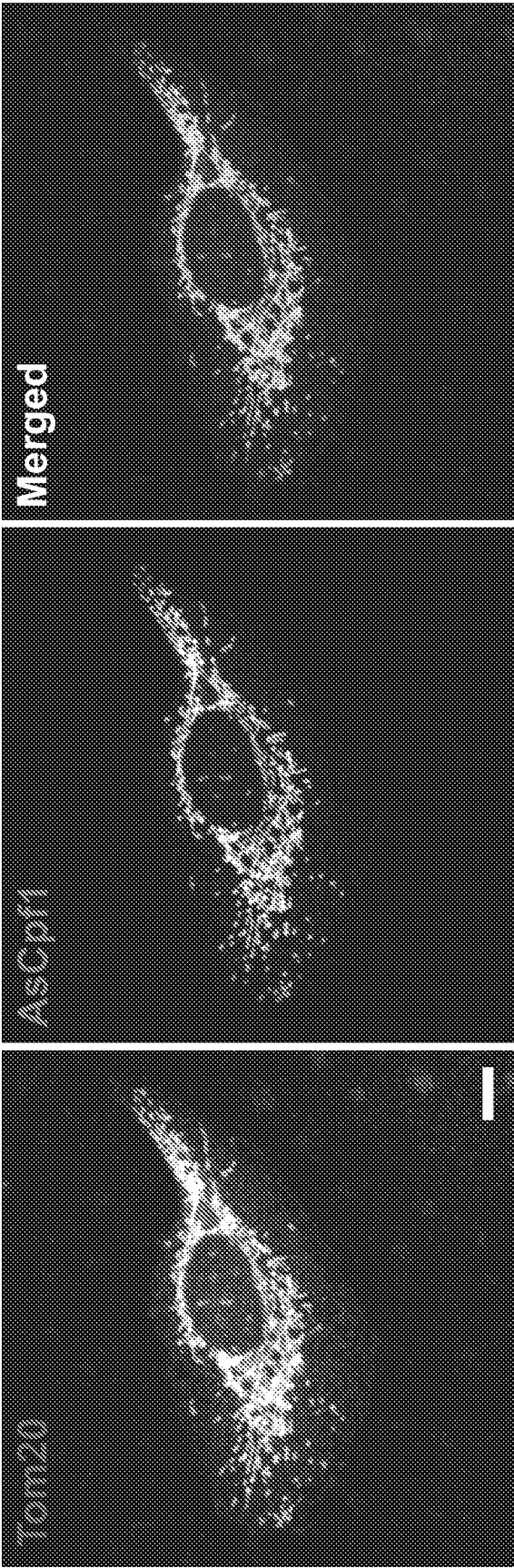
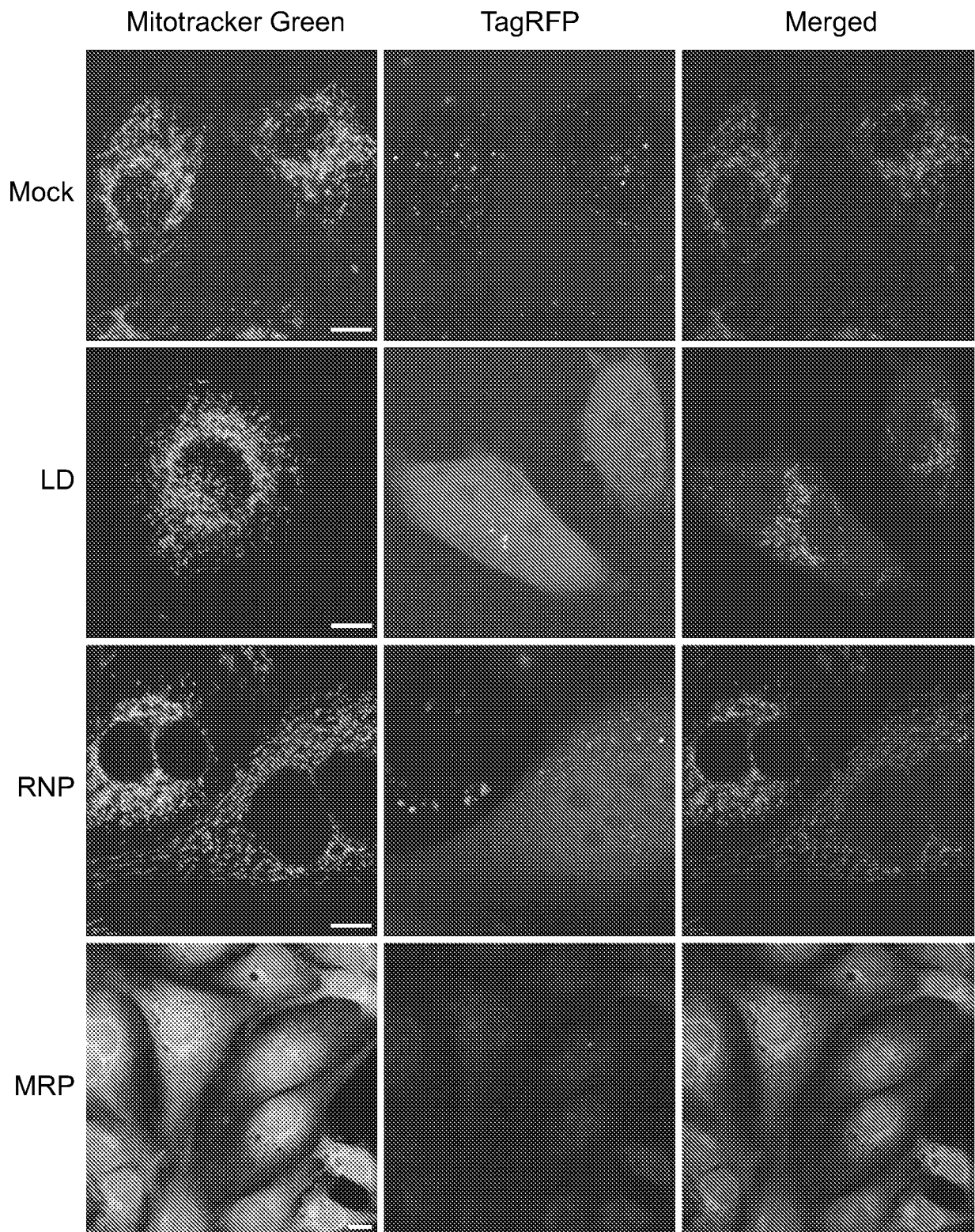


FIG. 6C

11/15

FIG. 7



12/15

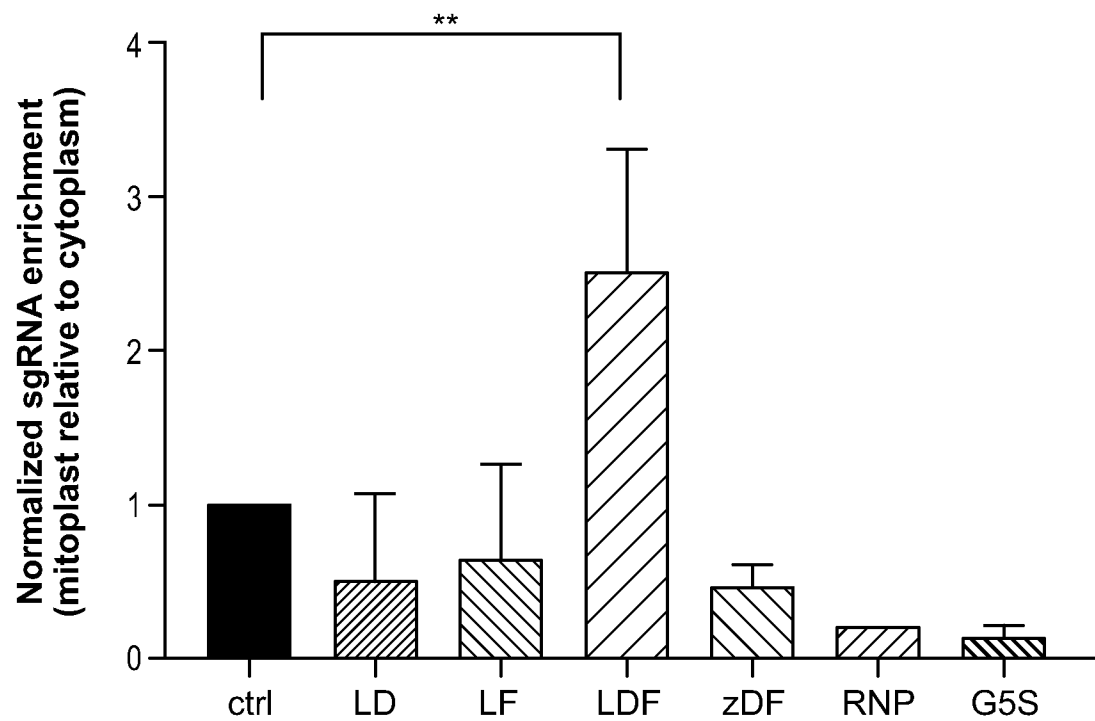


FIG. 8

13/15

FIG. 9A

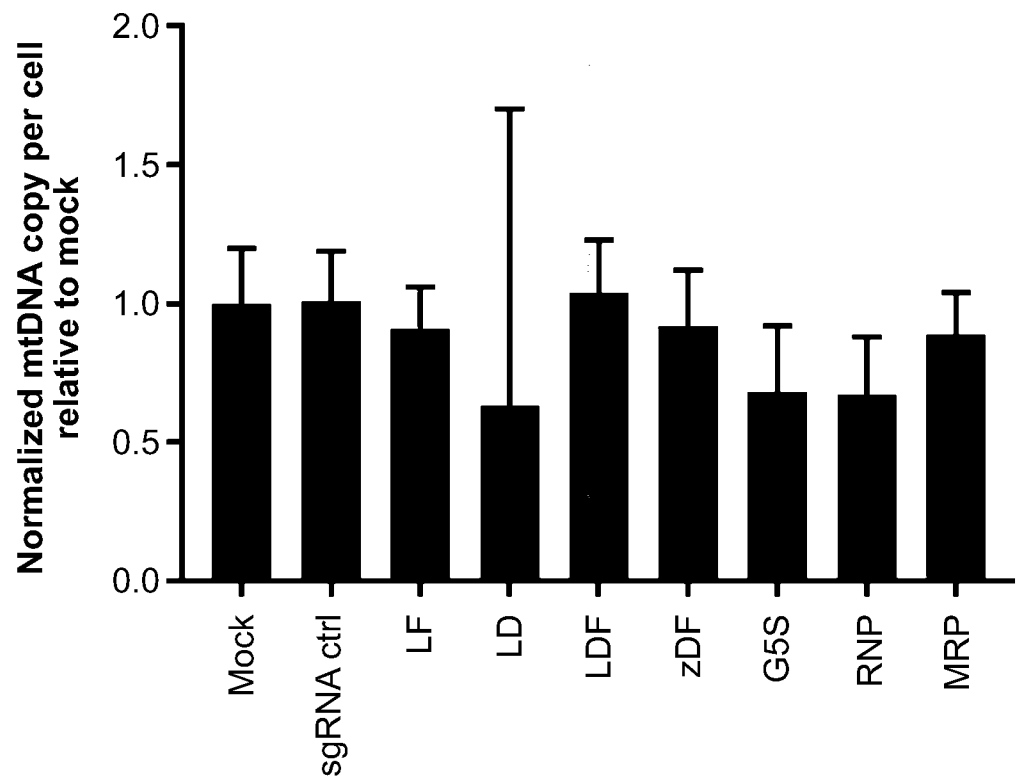


FIG. 9B

	CytB/Bactin	
	Mean	SD
Mock	1.00	0.20
sgRNA ctrl	1.01	0.18
LF	0.91	0.15
LD	0.63	1.07
LDF	1.04	0.19
zDF	0.92	0.20
G5S	0.68	0.24
RNP	0.67	0.21
MRP	0.89	0.15

14/15

FIG. 10A

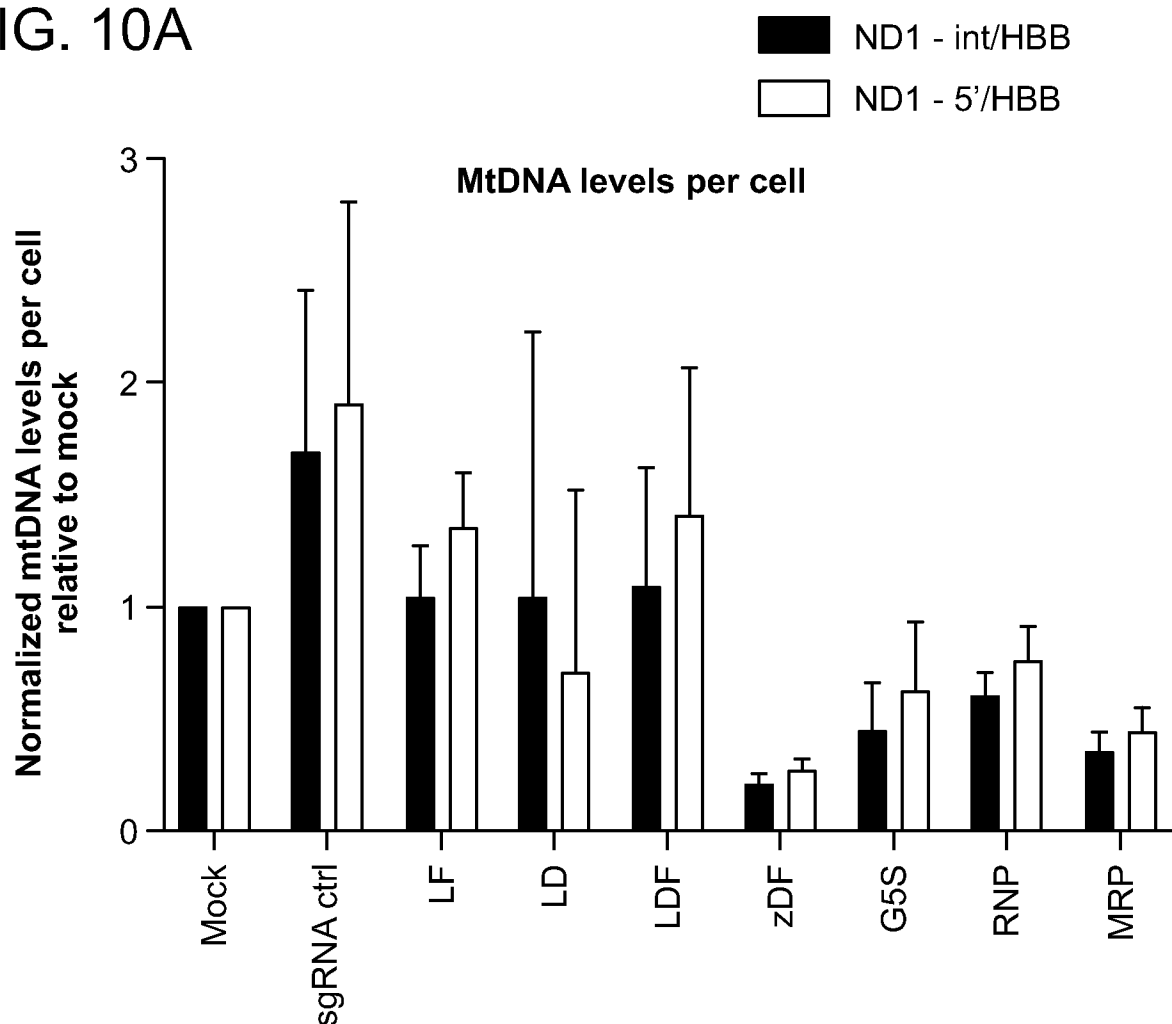


FIG. 10B

	ND1-int/HBB			ND1-5'/HBB		
	Mean	Upper Limit	Lower Limit	Mean	Upper Limit	Lower Limit
sgRNA ctrl	1.69	1.18	2.41	1.91	1.29	2.81
LF	1.04	0.85	1.27	1.35	1.15	1.60
LD	1.04	0.49	2.23	0.71	0.33	1.52
LDF	1.09	0.73	1.62	1.41	0.96	2.07
zDF	0.21	0.17	0.25	0.27	0.22	0.32
G5S	0.45	0.30	0.66	0.62	0.41	0.93
RNP	0.60	0.51	0.70	0.76	0.62	0.91
MRP	0.35	0.29	0.44	0.44	0.35	0.55

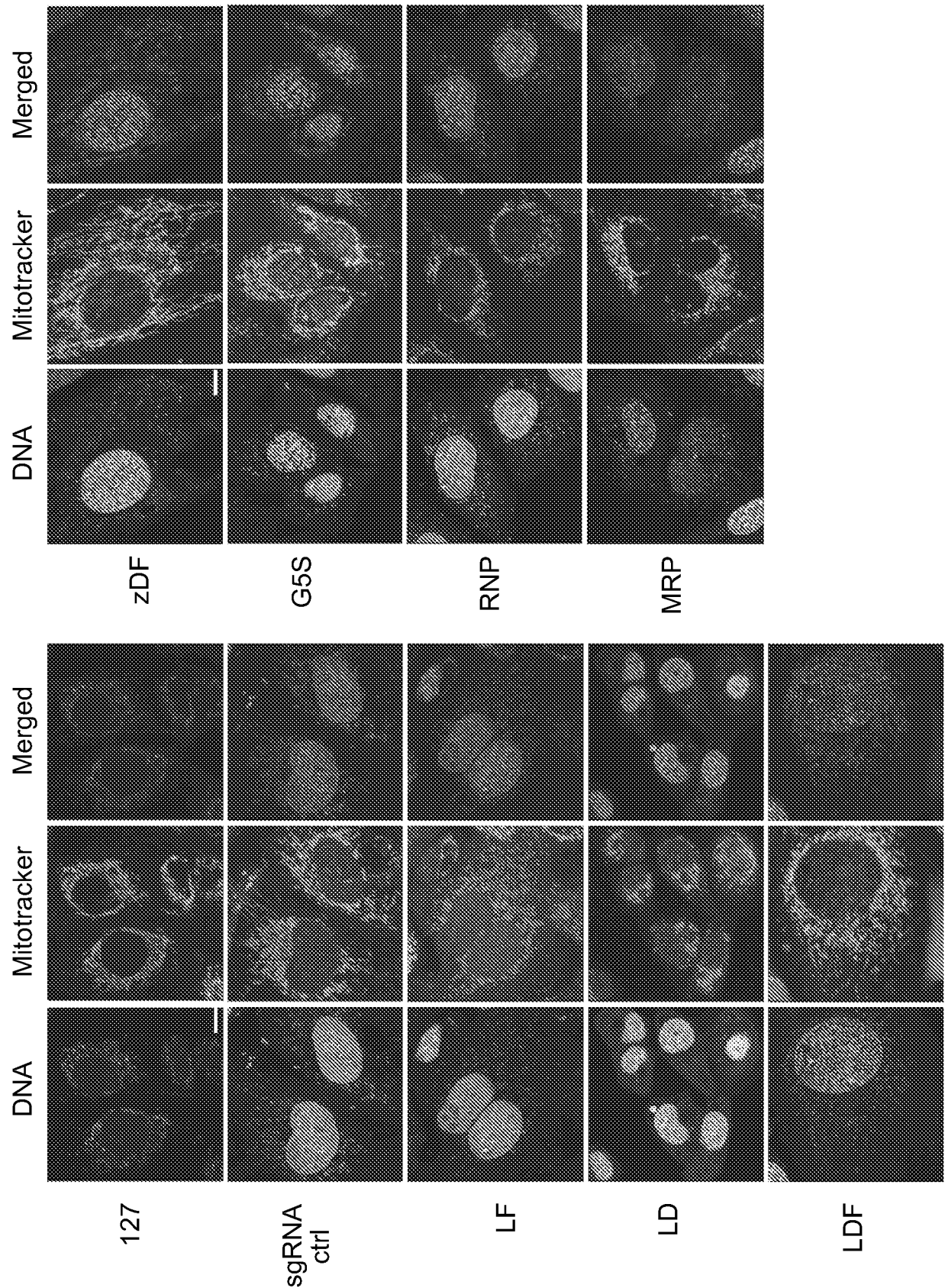


FIG. 11

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/036821

Box No. 1 Nucleotide and/or amino acid sequence(s) (Continuation of item 1.c of the first sheet)

1. With regard to any nucleotide and/or amino acid sequence disclosed in the international application, the international search was carried out on the basis of a sequence listing:
 - a. ☒ forming part of the international application as filed:
 - ☒ in the form of an Annex C/ST.25 text file.
 - ☐ on paper or in the form of an image file.
 - b. ☐ furnished together with the international application under PCT Rule 13ter.1(a) for the purposes of international search only in the form of an Annex C/ST.25 text file.
 - c. ☐ furnished subsequent to the international filing date for the purposes of international search only:
 - ☐ in the form of an Annex C/ST.25 text file (Rule 13ter.1(a)).
 - ☐ on paper or in the form of an image file (Rule 13ter.1(b) and Administrative Instructions, Section 713).
2. ☐ In addition, in the case that more than one version or copy of a sequence listing has been filed or furnished, the required statements that the information in the subsequent or additional copies is identical to that forming part of the application as filed or does not go beyond the application as filed, as appropriate, were furnished.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/036821

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☒ Claims Nos.: 26-29, 35, 37
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☐ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.
- ☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.
- ☐ No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/036821

A. CLASSIFICATION OF SUBJECT MATTER

IPC(8) - C12N 9/22; C12N 15/09; C12N 15/63; C12N 15/85 (2017.01)

CPC - C12N 9/22; C12N 15/102; C12N 15/907; C12N 2310/20 (2017.08)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

See Search History document

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

USPC - 435/455; 435/458; 435/463; 435/478 (keyword delimited)

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

See Search History document

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2015/071474 A2 (CRISPR THERAPEUTICS AG) 21 May 2015 (21.05.2015) entire document	1-4, 6, 8-12, 14-19, 21-25, 30-34, 38, 39, 41, 43, 44
Y		5, 7, 13, 20, 36, 40, 42
Y	CN 105602993 A (SHANGHAI SAI INK BIOLOGICAL TECHNOLOGY CO LTD) 25 May 2016 (25.05.2016) entire document; see machine translation	5, 40
Y	WO 2016/073433 A1 (E. I. DU PONT DE NEMOURS AND COMPANY) 12 May 2016 (12.05.2016) entire document	7, 42
Y	US 2015/0361450 A1 (CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE CNRS et al) 17 December 2015 (17.12.2015) entire document	13
Y	CN 105602935 A (NIE LING YUN) 25 May 2016 (25.05.2016) entire document; see machine translation	20
Y	WO 2014/085830 A2 (THE PARKINSON'S INSTITUTE et al) 05 June 2014 (05.06.2014) entire document	36
A	WO 2016/073990 A2 (EDITAS MEDICINE, INC.) 12 May 2016 (12.05.2016) entire document	1-25, 30-34, 36, 38-44
A	US 2015/0259684 A1 (HARVARD COLLEGE) 17 September 2015 (17.09.2015) entire document	1-25, 30-34, 36, 38-44



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T"

later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X"

document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y"

document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"

document member of the same patent family

Date of the actual completion of the international search

01 September 2017

Date of mailing of the international search report

25 SEP 2017

Name and mailing address of the ISA/US

Mail Stop PCT, Attn: ISA/US, Commissioner for Patents

P.O. Box 1450, Alexandria, VA 22313-1450

Facsimile No. 571-273-8300

Authorized officer

Blaine R. Copenheaver

PCT Helpdesk: 571-272-4300

PCT OSP: 571-272-7774

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US2017/036821

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JO et al. "Efficient Mitochondrial Genome Editing by CRISPR/Cas9," Biomed Research International, 10 September 2015 (10.09.2015), Vol. 2015, Article ID 305716, Pgs. 1-10. entire document	1-25, 30-34, 36, 38-44
A	US 9,139,628 B2 (MINCZUK et al) 22 September 2015 (22.09.2015) entire document	1-25, 30-34, 36, 38-44
A	US 2016/0138008 A1 (UNIVERSITY WIEN et al) 19 May 2016 (19.05.2016) entire document	1-25, 30-34, 36, 38-44