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

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
14 May 2020 (14.05.2020)

WIPO | PCT

(10) International Publication Number
WO 2020/097445 A1

(51) International Patent Classification:
C12N 9/22 (2006.01) *C07K 19/00* (2006.01)
C12N 15/81 (2006.01) *C12N 15/90* (2006.01)
C12N 15/86 (2006.01) *C07K 14/195* (2006.01)

(21) International Application Number:
PCT/US2019/060442

(22) International Filing Date:
08 November 2019 (08.11.2019)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
62/758,307 09 November 2018 (09.11.2018) US
62/832,130 10 April 2019 (10.04.2019) US

(71) Applicant: **INARI AGRICULTURE, INC.** [US/US]; One Kendall Square, Building 600/700, Suite 7-501, Cambridge, Massachusetts 02139 (US).

(72) Inventors: **PYZOCHA, Neena Kenton**; c/o Inari Agriculture, Inc., One Kendall Square, Building 600/700, Suite 7-501, Cambridge, Massachusetts 02139 (US). **JOYCE, Adam Patrick**; c/o Inari Agriculture, Inc., One Kendall Square, Building 600/700, Suite 7-501, Cambridge, Massachusetts 02139 (US). **KREMLING, Karl Anton Grothe**; c/o Inari Agriculture, Inc., One Kendall Square, Building 600/700, Suite 7-501, Cambridge, Massachusetts 02139 (US).

(74) Agent: **ROMANO, Charles P.** et al.; Thompson Coburn LLP, One US Bank Plaza, St. Louis, Missouri 63101 (US).

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

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

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with sequence listing part of description (Rule 5.2(a))*

(54) Title: RNA-GUIDED NUCLEASES AND DNA BINDING PROTEINS

(57) Abstract: Compositions and methods related to Cas proteins, nucleic acids encoding the Cas proteins, and modified host cells comprising the Cas proteins and/or encoding nucleic acids are disclosed. Cas proteins are useful in a variety of applications. Cas proteins bind guide RNAs that in turn provide functional specificity to the Cas proteins, nucleic acids encoding the Cas guide RNAs, and modified host cells comprising the Cas guide RNAs and/or encoding nucleic acids. The Cas polypeptides and corresponding guide RNAs can be used in a variety of applications.

WO 2020/097445 A1

RNA-Guided Nucleases and DNA Binding Proteins

Reference to Related Applications

[0001] This application claims benefit of priority to U.S. Provisional Patent Application No. 62/758307 filed 9 November 2018, and to U.S. Provisional Patent Application No. 62/832,130 filed 10 April 2019,which are each incorporated herein by reference in their entireties.

Incorporation of Sequence Listing

[0002] A sequence listing containing the file named "10020WO_ST25.txt" which is 42116 bytes (measured in MS-Windows®), comprises 78 biological sequences, and was created on November7, 2019, is electronically filed herewith via the USPTO's EFS system, and is incorporated herein by reference in its entirety.

Field

[0003] The invention is generally related to CRISPR effector systems.

Background

[0004] The CRISPR/Cas system of bacterial acquired immunity against phages and viruses has been adapted into potent new technologies for genomic modifications, as well as other research tools. A few Class 2 nucleases have been intensively used and characterized, yet a need remains for alternative nucleases with different properties that may provide optimal performance or options in a variety genome modification or diagnostic applications.

Summary

[0005] The present disclosure provides RNA-guided endonuclease polypeptides, referred to herein as "CasP6" polypeptides (also referred to as "CasP6 proteins"); nucleic acids encoding the CasP6 polypeptides; and modified host cells comprising the CasP6 polypeptides and/or nucleic acids encoding same. CasP6 polypeptides are useful in a variety of applications, which are provided.

[0006] The present disclosure provides guide RNAs (referred to herein as "CasP6 guide RNAs") that bind to and provide sequence specificity to the CasP6 proteins; nucleic acids encoding the CasP6 guide RNAs; and modified host cells comprising the CasP6 guide RNAs and/or nucleic acids encoding same. CasP6 guide RNAs are useful in a variety of applications, which are provided.

[0007] Also provided are nucleic acid detection systems and methods that comprise the CasP6 polypeptides and guide RNAs. In certain embodiments, nucleic acid detection systems comprising: (i) a CasP6 polypeptide or CasP6 fusion polypeptide; (ii) one or more CasP6 guide RNAs each comprising a guide sequence that can hybridize to RNA molecules having corresponding target sequences; and (iii) an RNA-based masking construct are provided. In

certain embodiments, the RNA-based masking construct does not comprise the corresponding target sequence(s). Methods of detecting a target RNA molecule comprising: (a) contacting the target RNA molecule with: (i) a CasP6 polypeptide or a CasP6 fusion polypeptide; (ii) one or more CasP6 guide RNAs each comprising a guide sequence that can hybridize to corresponding target sequences in the target RNA molecule; and (iii) an RNA-based masking construct; wherein the CasP6 comprises endonuclease enzyme activity and wherein the RNA-based masking construct does not comprise the target sequence; and (b) assaying for modification of the RNA-based masking construct.

Brief Description of the Drawings

[0008] Figure 1 depicts three CasP6 protein sequences: CasP6A (SEQ ID NO: 1), CasP6B (SEQ ID NO: 2), and CasP6C (SEQ ID NO: 3).

[0009] Figure 2 diagrams the RuvC domain location within the CasP6A (top), CasP6B (middle), and CasP6C (bottom) primary structure.

[0010] Figure 3A,B,C,D shows the direct repeat (DR) DNA sequences corresponding to CasP6A (SEQ ID NO: 4), CasP6B (SEQ ID NO: 5), and CasP6C (SEQ ID NO: 6) as well as the predicted full length RNA structures encoded by the DNA sequences (SEQ ID NO: 7, SEQ ID NO: 8, and SEQ ID NO:9, respectively).

[0011] Figure 4 A and B show the temperature dependence of CasP6B RNA cleavage.

Detailed Description

Definitions

[0012] The term "and/or" where used herein is to be taken as specific disclosure of each of the two or more specified features or components with or without the other specified features. Thus, the term "and/or" as used in a phrase such as "A and/or B" herein is intended to include "A and B," "A or B," "A" (alone), and "B" (alone). Likewise, the term "and/or" as used in a phrase such as "A, B, and/or C" is intended to encompass each of the following embodiments: A, B, and C; A, B, or C; A or C; A or B; B or C; A and C; A and B; B and C; A (alone); B (alone); and C (alone).

[0013] As used herein, the terms “correspond,” “corresponding,” and the like, when used in the context of an amino acid position, mutation, and/or substitution in any given CasP6 polypeptide with respect to the reference CasP6 polypeptide sequence of any one of SEQ ID NO: 1-3, all refer to the position, mutation, and/or substitution of the amino acid residue in the given CasP6 sequence that has identity or similarity to the amino acid residue in the reference CasP6 polypeptide sequence of SEQ ID NO: 1-3 when the given CasP6 polypeptide is aligned to the

reference CasP6 polypeptide sequence of SEQ ID NO: 1-3 using a pairwise alignment algorithm (*e.g.* CLUSTAL O 1.2.4 with default parameters).

[0014] As used herein, the phrase "DNA donor template" refers to a DNA molecule having homology to the target editing site. DNA donor template molecules can be used to edit a target editing site in a genome by homology-directed repair.

[0015] As used herein, the terms "include," "includes," and "including" are to be construed as at least having the features to which they refer while not excluding any additional unspecified features.

[0016] "Heterologous," as used herein, means a nucleotide or polypeptide sequence that is not found in the native nucleic acid or protein, respectively. For example, relative to a CasP6 polypeptide, a heterologous polypeptide comprises an amino acid sequence from a protein other than the CasP6 polypeptide. In some cases, a portion of a CasP6 protein from one species is fused to a portion of a Cas protein from a different species. The Cas sequence from each species could therefore be considered to be heterologous relative to one another. As another example, a CasP6 protein (e.g., a dCasP6 protein) can be fused to an active domain from a non-CasP6 protein (e.g., a histone deacetylase), and the sequence of the active domain could be considered a heterologous polypeptide (it is heterologous to the CasP6 protein).

[0017] The terms "polynucleotide" and "nucleic acid," used interchangeably herein, refer to a polymeric form of nucleotides of any length, either ribonucleotides or deoxynucleotides. Thus, this term includes, but is not limited to, single-, double-, or multi-stranded DNA or RNA, genomic DNA, cDNA, DNA-RNA hybrids, or a polymer comprising purine and pyrimidine bases or other natural, chemically or biochemically modified, non-natural, or derivatized nucleotide bases. The terms "polynucleotide" and "nucleic acid" should be understood to include, as applicable to the embodiment being described, single- stranded (such as sense or antisense) and double-stranded polynucleotides.

[0018] The terms "polypeptide," "peptide," and "protein", are used interchangeably herein, refer to a polymeric form of amino acids of any length, which can include genetically coded and non-genetically coded amino acids, chemically or biochemically modified or derivatized amino acids, and polypeptides having modified peptide backbones. The term includes fusion proteins, including, but not limited to, fusion proteins with a heterologous amino acid sequence, fusions with heterologous and homologous leader sequences, with or without N-terminal methionine residues; immunologically tagged proteins; and the like.

[0019] The term "naturally-occurring" as used herein as applied to a nucleic acid, a protein, a cell, or an organism, refers to a nucleic acid, cell, protein, or organism that is found in nature.

[0020] As used herein the term "isolated" is meant to describe a polynucleotide, a polypeptide, or a cell that is in an environment different from that in which the polynucleotide, the polypeptide, or the cell naturally occurs. An isolated genetically modified host cell may be present in a mixed population of genetically modified host cells.

[0021] As used herein, the term "exogenous nucleic acid" refers to a nucleic acid that is not normally or naturally found in and/or produced by a given bacterium, organism, or cell in nature. As used herein, the term "endogenous nucleic acid" refers to a nucleic acid that is normally found in and/or produced by a given bacterium, organism, or cell in nature. An "endogenous nucleic acid" is also referred to as a "native nucleic acid" or a nucleic acid that is "native" to a given bacterium, organism, or cell.

[0022] "Recombinant," as used herein, means that a particular nucleic acid (DNA or RNA) is the product of various combinations of cloning, restriction, and/or ligation steps resulting in a construct having a structural coding or non-coding sequence distinguishable from endogenous nucleic acids found in natural systems. Generally, DNA sequences encoding the structural coding sequence can be assembled from cDNA fragments and short oligonucleotide linkers, or from a series of synthetic oligonucleotides, to provide a synthetic nucleic acid which is capable of being expressed from a recombinant transcriptional unit contained in a cell or in a cell-free transcription and translation system. Such sequences can be provided in the form of an open reading frame uninterrupted by internal non-translated sequences, or introns, which are typically present in eukaryotic genes. Genomic DNA comprising the relevant sequences can also be used in the formation of a recombinant gene or transcriptional unit. Sequences of non-translated DNA may be present 5' or 3' from the open reading frame, where such sequences do not interfere with manipulation or expression of the coding regions, and may indeed act to modulate production of a desired product by various mechanisms (see "DNA regulatory sequences", below).

[0023] Thus, e.g., the term "recombinant" polynucleotide or "recombinant" nucleic acid refers to one which is not naturally occurring, e.g., is made by the artificial combination of two otherwise separated segments of sequence through human intervention. This artificial combination is often accomplished by either chemical synthesis means, or by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques. Such is usually done to replace a codon with a redundant codon encoding the same or a conservative amino acid, while typically introducing or removing a sequence recognition site. Alternatively, it is performed to join together nucleic acid segments of desired functions to generate a desired combination of functions. This artificial combination is often accomplished by either chemical synthesis means,

or by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques.

[0024] Similarly, the term "recombinant" polypeptide refers to a polypeptide which is not naturally occurring, e.g., is made by the artificial combination of two otherwise separated segments of amino sequence through human intervention. Thus, e.g., a polypeptide that comprises a heterologous amino acid sequence is recombinant.

[0025] By "construct" or "vector" is meant a recombinant nucleic acid, generally recombinant DNA, which has been generated for the purpose of the expression and/or propagation of a specific nucleotide sequence(s), or is to be used in the construction of other recombinant nucleotide sequences.

[0026] The terms "DNA regulatory sequences," "control elements," and "regulatory elements," used interchangeably herein, refer to transcriptional and translational control sequences, such as promoters, enhancers, polyadenylation signals, terminators, protein degradation signals, and the like, that provide for and/or regulate expression of a coding sequence and/or production of an encoded polypeptide in a host cell.

[0027] The term "transformation" is used interchangeably herein with "genetic modification" and refers to a permanent or transient genetic change induced in a cell following introduction of new nucleic acid (e.g., DNA exogenous to the cell) into the cell. Genetic change ("modification") can be accomplished either by incorporation of the new nucleic acid into the genome of the host cell, or by transient or stable maintenance of the new nucleic acid as an episomal element. Where the cell is a eukaryotic cell, a permanent genetic change is generally achieved by introduction of new DNA into the genome of the cell. In prokaryotic cells, permanent changes can be introduced into the chromosome or via extrachromosomal elements such as plasmids and expression vectors, which may contain one or more selectable markers to aid in their maintenance in the recombinant host cell. Suitable methods of genetic modification include viral infection, transfection, conjugation, protoplast fusion, electroporation, particle gun technology, calcium phosphate precipitation, direct microinjection, and the like. The choice of method is generally dependent on the type of cell being transformed and the circumstances under which the transformation is taking place (i.e. in vitro, ex vivo, or in vivo). A general discussion of these methods can be found in Ausubel, et al, Short Protocols in Molecular Biology, 3rd ed., Wiley & Sons, 1995.

[0028] "Operably linked" refers to a juxtaposition wherein the components so described are in a relationship permitting them to function in their intended manner. For instance, a promoter is operably linked to a coding sequence if the promoter affects its transcription or expression. As used herein, the terms "heterologous promoter" and "heterologous control regions" refer to

promoters and other control regions that are not normally associated with a particular nucleic acid in nature. For example, a "transcriptional control region heterologous to a coding region" is a transcriptional control region that is not normally associated with the coding region in nature.

[0029] A "host cell," as used herein, denotes an in vivo or in vitro eukaryotic cell, a prokaryotic cell, or a cell from a multicellular organism (e.g., a cell line) cultured as a unicellular entity, which eukaryotic or prokaryotic cells can be, or have been, used as recipients for a nucleic acid (e.g., an expression vector), and include the progeny of the original cell which has been genetically modified by the nucleic acid. It is understood that the progeny of a single cell may not necessarily be completely identical in morphology or in genomic or total DNA complement as the original parent, due to natural, accidental, or deliberate mutation. A "recombinant host cell" (also referred to as a "genetically modified host cell") is a host cell into which has been introduced a heterologous nucleic acid, e.g., an expression vector. For example, a subject prokaryotic host cell is a genetically modified prokaryotic host cell (e.g., a bacterium), by virtue of introduction into a suitable prokaryotic host cell of a heterologous nucleic acid, e.g., an exogenous nucleic acid that is foreign to (not normally found in nature in) the prokaryotic host cell, or a recombinant nucleic acid that is not normally found in the prokaryotic host cell; and a subject eukaryotic host cell is a genetically modified eukaryotic host cell, by virtue of introduction into a suitable eukaryotic host cell of a heterologous nucleic acid, e.g., an exogenous nucleic acid that is foreign to the eukaryotic host cell, or a recombinant nucleic acid that is not normally found in the eukaryotic host cell.

[0030] The term "conservative amino acid substitution" refers to the interchangeability in proteins of amino acid residues having similar side chains. For example, a group of amino acids having aliphatic side chains consists of glycine, alanine, valine, leucine, and isoleucine; a group of amino acids having aliphatic-hydroxyl side chains consists of serine and threonine; a group of amino acids having amide-containing side chains consists of asparagine and glutamine; a group of amino acids having aromatic side chains consists of phenylalanine, tyrosine, and tryptophan; a group of amino acids having basic side chains consists of lysine, arginine, and histidine; and a group of amino acids having sulfur- containing side chains consists of cysteine and methionine. Exemplary conservative amino acid substitution groups are: valine -leucine -isoleucine, phenylalanine -tyrosine, lysine-arginine, alanine-valine, and asparagine-glutamine.

[0031] A polynucleotide or polypeptide has a certain percent "sequence identity" to another polynucleotide or polypeptide, meaning that, when aligned, that percentage of bases or amino acids are the same, and in the same relative position, when comparing the two sequences. Sequence similarity can be determined in a number of different manners. To determine sequence

identity, sequences can be aligned using the methods and computer programs, including BLAST, available over the world wide web at ncbi.nlm.nih.gov/BLAST. See, e.g., Altschul et al. (1990), /. Mol. Biol. 215:403-10. Another alignment algorithm is FASTA, available in the Genetics Computing Group (GCG) package, from Madison, Wisconsin, USA, a wholly owned subsidiary of Oxford Molecular Group, Inc. Other techniques for alignment are described in Methods in Enzymology, vol. 266: Computer Methods for Macromolecular Sequence Analysis (1996), ed. Doolittle, Academic Press, Inc., a division of Harcourt Brace & Co., San Diego, California, USA. Of particular interest are alignment programs that permit gaps in the sequence. The Smith-Waterman is one type of algorithm that permits gaps in sequence alignments. See Meth. Mol. Biol. 70: 173-187 (1997). Also, the GAP program using the Needleman and Wunsch alignment method can be utilized to align sequences. See /. Mol. Biol. 48: 443-453 (1970).

[0032] As used herein, the terms "treatment," "treating," and the like, refer to obtaining a desired pharmacologic and/or physiologic effect. The effect may be prophylactic in terms of completely or partially preventing a disease or symptom thereof and/or may be therapeutic in terms of a partial or complete cure for a disease and/or adverse effect attributable to the disease. "Treatment," as used herein, covers any treatment of a disease in a mammal, e.g., in a human, and includes: (a) preventing the disease from occurring in a subject which may be predisposed to the disease but has not yet been diagnosed as having it; (b) inhibiting the disease, i.e., arresting its development; and (c) relieving the disease, i.e., causing regression of the disease.

[0033] The terms "individual," "subject," "host," and "patient," used interchangeably herein, refer to an individual organism, e.g., a mammal, including, but not limited to, murines, simians, humans, mammalian farm animals, mammalian sport animals, and mammalian pets.

[0034] Before the present invention is further described, it is to be understood that this invention is not limited to particular embodiments described, as such may, of course, vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to be limiting, since the scope of the present invention will be limited only by the appended claims.

[0035] Where a range of values is provided, it is understood that each intervening value, to the tenth of the unit of the lower limit unless the context clearly dictates otherwise, between the upper and lower limit of that range and any other stated or intervening value in that stated range, is encompassed within the invention. The upper and lower limits of these smaller ranges may independently be included in the smaller ranges, and are also encompassed within the invention, subject to any specifically excluded limit in the stated range. Where the stated range includes one

or both of the limits, ranges excluding either or both of those included limits are also included in the invention.

[0036] Unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials similar or equivalent to those described herein can also be used in the practice or testing of the present invention, the preferred methods and materials are now described. All publications mentioned herein are incorporated herein by reference to disclose and describe the methods and/or materials in connection with which the publications are cited.

[0037] It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural referents unless the context clearly dictates otherwise. Thus, for example, reference to "a CasP6 polypeptide" includes a plurality of such polypeptides and reference to "the guide RNA" includes reference to one or more guide RNAs and equivalents thereof known to those skilled in the art, and so forth. It is further noted that the claims may be drafted to exclude any optional element. As such, this statement is intended to serve as antecedent basis for use of such exclusive terminology as "solely," "only" and the like in connection with the recitation of claim elements, or use of a "negative" limitation.

[0038] It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable sub-combination. All combinations of the embodiments pertaining to the invention are specifically embraced by the present invention and are disclosed herein just as if each and every combination was individually and explicitly disclosed. In addition, all sub-combinations of the various embodiments and elements thereof are also specifically embraced by the present invention and are disclosed herein just as if each and every such subcombination was individually and explicitly disclosed herein.

[0039] To the extent to which any of the preceding definitions is inconsistent with definitions provided in any patent or non-patent reference incorporated herein by reference, any patent or non-patent reference cited herein, or in any patent or non-patent reference found elsewhere, it is understood that the preceding definition will be used herein.

Further Description

[0040] The present disclosure provides RNA-guided endonuclease polypeptides, referred to herein as "CasP6" polypeptides (also referred to as "CasP6 proteins"); nucleic acids encoding the

CasP6 polypeptides; and modified host cells comprising the CasP6 polypeptides and/or nucleic acids encoding same. CasP6 polypeptides are useful in a variety of applications, which are provided.

[0041] The present disclosure provides guide RNAs (referred to herein as "CasP6 guide RNAs") that bind to and provide sequence specificity to the CasP6 proteins; nucleic acids encoding the CasP6 guide RNAs; and modified host cells comprising the CasP6 guide RNAs and/or nucleic acids encoding same. CasP6 guide RNAs are useful in a variety of applications, which are provided.

[0042] A CRISPR/Cas endonuclease (e.g., a CasP6 protein) interacts with (binds to) a corresponding guide RNA (e.g., a CasP6 guide RNA) to form a ribonucleoprotein (RNP) complex that is targeted to a particular site in a target nucleic acid via base pairing between the guide RNA and a target sequence within the target nucleic acid molecule. A guide RNA includes a nucleotide sequence (a guide sequence) that is complementary to a sequence (the target site) of a target nucleic acid. Thus, a CasP6 protein forms a complex with a CasP6 guide RNA and the guide RNA provides sequence specificity to the RNP complex via the guide sequence. The CasP6 protein of the complex provides the site-specific activity. In other words, the CasP6 protein is guided to a target site (e.g., stabilized at a target site) within a target nucleic acid sequence (e.g. a chromosomal sequence or an extrachromosomal sequence, e.g., an episomal sequence, a minicircle sequence, a mitochondrial sequence, a chloroplast sequence, etc.) by virtue of its association with the guide RNA.

[0043] The present disclosure provides compositions comprising a CasP6 polypeptide (and/or a nucleic acid encoding the CasP6 polypeptide) (e.g., where the CasP6 polypeptide can be a naturally existing protein, a nickase CasP6 protein, a dCasP6 protein, a chimeric CasP6 protein, etc.). The present disclosure provides compositions comprising a CasP6 guide RNA (and/or a nucleic acid encoding the CasP6 guide RNA) (e.g., where the CasP6 guide RNA can be in single guide format). The present disclosure provides compositions comprising (a) a CasP6 polypeptide (and/or a nucleic acid encoding the CasP6 polypeptide) (e.g., where the CasP6 polypeptide can be a naturally existing protein, a nickase CasP6 protein, a dCasP6 protein, a chimeric CasP6 protein, etc.) and (b) a CasP6 guide RNA (and/or a nucleic acid encoding the CasP6 guide RNA) (e.g., where the CasP6 guide RNA can be in single guide format). The present disclosure provides a nucleic acid/protein complex (RNP complex) comprising: (a) a CasP6 polypeptide of the present disclosure (e.g., where the CasP6 polypeptide can be a naturally existing protein, a nickase CasP6 protein, a dCasP6 protein, a chimeric CasP6 protein, etc.); and (b) a CasP6 guide RNA (e.g., where the CasP6 guide RNA can be in single guide format).

[0044] A CasP6 polypeptide (this term is used interchangeably with the term "CasP6 protein") can bind and/or modify (e.g., cleave, nick, methylate, demethylate, etc.) a target nucleic acid and/or a polypeptide associated with target nucleic acid (e.g., methylation or acetylation of a histone tail) (e.g., in some cases the CasP6 protein includes a fusion partner with an activity, and in some cases the CasP6 protein provides nuclease activity). In some cases, the CasP6 protein is a naturally-occurring protein (e.g., naturally occurs in prokaryotic cells). In other cases, the CasP6 protein is not a naturally-occurring polypeptide (e.g., the CasP6 protein is a variant CasP6 protein, a chimeric protein, and the like).

[0045] Assays to determine whether given protein interacts with a CasP6 guide RNA can be any convenient binding assay that tests for binding between a protein and a nucleic acid. Suitable binding assays (e.g., gel shift assays) will be known to one of ordinary skill in the art (e.g., assays that include adding a CasP6 guide RNA and a protein to a target nucleic acid). Assays to determine whether a protein has an activity (e.g., to determine if the protein has nuclease activity that cleaves a target nucleic acid and/or some heterologous activity) can be any convenient assay (e.g., any convenient nucleic acid cleavage assay that tests for nucleic acid cleavage). Suitable assays (e.g., cleavage assays) will be known to one of ordinary skill in the art.

[0046] A naturally occurring CasP6 protein functions as an endonuclease that catalyzes a strand break (double or single strand) at a specific sequence in a targeted DNA or RNA. The sequence specificity is provided by the associated guide RNA, which hybridizes to a target sequence within the target DNA or RNA. The naturally occurring guide RNA may include a tracrRNA hybridized to a crRNA, where the crRNA includes a guide sequence that hybridizes to a target sequence in the target DNA or RNA.

[0047] As used herein, CasP6 endonuclease activity refers to CRISPR endonuclease activity wherein, a guide RNA associated with a CasP6 polypeptide causes the CasP6-guide RNA complex to bind to a pre-determined nucleotide sequence that is complementary to the gRNA; and wherein CasP6 endonuclease activity can introduce a strand break at or near the site targeted by the gRNA. In certain embodiments, this this is a double-stranded break, and it may be a blunt or a staggered DNA double-stranded break. As used herein a "staggered DNA double-stranded break" can result in a double strand break with about 1, about 2, about 3, about 4, about 5, about 6, about 7, about 8, about 9, or about 10 nucleotides of overhang on either the 3' or 5' ends following cleavage. The double strand break can occur at or near the sequence to which the guide sequence is targeted. In some embodiments, the CasP6 protein of the subject methods and/or compositions is (or is derived from) a naturally occurring (wild type) protein. Examples of

naturally occurring CasP6 proteins are depicted in Figure 1 and are set forth as SEQ ID NOS: 1-3.

[0048] In some cases, a CasP6 protein (of the subject compositions and/or methods) includes an amino acid sequence having 20% or more sequence identity (e.g., 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100% sequence identity) with the CasP6 protein sequence set forth as any of SEQ ID NOS: 1-3. For example, in some cases, a CasP6 protein includes an amino acid sequence having 50% or more sequence identity (e.g., 60% or more, 70% or more, 80% or more, 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100% sequence identity) with the CasP6 protein sequence set forth as any of SEQ ID NOS: 1-3. In some cases, a CasP6 protein includes an amino acid sequence having 80% or more sequence identity (e.g., 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100% sequence identity) with the CasP6 protein sequence set forth as any of SEQ ID NOS: 1-3. In some cases, a CasP6 protein includes an amino acid sequence having 90% or more sequence identity (e.g., 95% or more, 97% or more, 98% or more, 99% or more, or 100% sequence identity) with the CasP6 protein sequence set forth as any of SEQ ID NOS: 1-3. In some cases, a CasP6 protein includes an amino acid sequence having the CasP6 protein sequence set forth as any of SEQ ID NOS: 1-3. In some cases, a CasP6 protein includes an amino acid sequence having the CasP6 protein sequence set forth as any of SEQ ID NOS: 1-3, with the exception that the sequence includes an amino acid substitution (e.g., 1, 2, or 3 amino acid substitutions) that reduces the naturally occurring catalytic activity of the protein (e.g., such as at amino acid positions described below). In certain embodiments, the CasP6 protein can comprise one or more conservative amino acid substitutions in a CasP6 protein sequence set forth in any one of SEQ ID NOs: 1-3. In certain embodiments, the CasP3 protein can comprise one or more amino acid substitutions in a CasP6 protein sequence set forth in any one of SEQ ID NOs: 1-3 where the substituted amino acid is a corresponding amino acid in a distinct CasP6 protein sequence set forth in any one of SEQ ID NOs: 1-3.

[0049] The domains of a CasP6 protein are depicted in Figure 2. As can be seen in the schematic representation of Figure 2 (amino acids are numbered based on the CasP6 proteins, SEQ ID NOs: 1-3), a CasP6 protein has a C-terminal domain that includes 3 partial RuvC domains (RuvC-I, RuvC-II, and RuvC-III, also referred to herein as subdomains) that are not contiguous with respect to the primary amino acid sequence of the CasP6 protein, but form a RuvC domain once the protein is produced and folds.

[0050] Thus, in some cases, a CasP6 protein (of the subject compositions and/or methods) includes an amino acid sequence having a length in a range of from 650-950 amino acids (e.g, from 700-900, 700-850, 750-900, 750-850, 700-800, 800-900, 700-750, 750-800, 800-850, 850-900, 700-725, 725-750, 750-775, 775-800, 800-825, 825-850, 850-875, 875-900, or 731-864 amino acids). In some cases, a CasP6 protein (of the subject compositions and/or methods) includes an amino acid sequence having 20% or more sequence identity (e.g., 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100% sequence identity) with any of the CasP6 proteins set forth as SEQ ID NOs: 1-3.

[0051] In some cases, a CasP6 protein (of the subject compositions and/or methods) includes an amino acid sequence, that includes a split RuvC domain (e.g., 3 partial RuvC domains RuvC-I, RuvC- II, and RuvC-III). In some cases, a CasP6 protein (of the subject compositions and/or methods) includes a split RuvC domain (e.g., 3 partial RuvC domains - RuvC-I, RuvC-II, and RuvC-III) with amino acid sequence having 20% or more sequence identity (e.g., 30% or more, 40% or more, 50% or more, 60% or more, 70% or more, 80% or more, 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100% sequence identity) with the split RuvC domain of any one of SEQ ID NOs: 1-3. For example, in some cases, a CasP6 protein includes an amino acid sequence having 50% or more sequence identity (e.g., 60% or more, 70% or more, 80% or more, 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100% sequence identity) with the split RuvC domain of SEQ ID NO: 1, SEQ ID NO: 2, OR SEQ ID NO: 3. In certain embodiments, an amino acid in an RuvC-I, RuvC-II, and RuvC-III subdomain of a CasP6 protein can be substituted with a corresponding amino acid in an RuvC-I, RuvC-II, and/or RuvC-III subdomain of SEQ ID NO:1, SEQ ID NO: 2, or SEQ ID NO: 3. In certain embodiments, the corresponding amino acid is an amino acid having similar side chain with respect to polarity and/or size. In many cases a CasP6 protein has the conserved residues of the RuvC subdomains and/or one or more of the corresponding residues in the RuvC-I, RuvC-II, or RuvC-III domain of a CasP6 protein of SEQ ID NO: 1-3.

[0052] The RuvC subdomains of SEQ ID NO: 1 include the following catalytic residues: RuvC-I – AGLDVGWR (SEQ ID NO: 10), RuvC-II – LSVEGSFL (SEQ ID NO: 11), and RuvC-III – QDHNASVNL (SEQ ID NO: 12). The RuvC subdomains of SEQ ID NO: 2 include the following catalytic residues: RuvC-I – AGLDIGWR (SEQ ID NO: 13), RuvC-II -- LIVETSFL (SEQ ID NO: 14), and RuvC-III – QDQNASVNL (SEQ ID NO: 15). The RuvC subdomains of SEQ ID NO: 3 include the following catalytic residues: RuvC-I – AGLDIGWY (SEQ ID NO:

16), RuvC-II – LGFEEDFL (SEQ ID NO: 17), and RuvC-III – QDENAAVNL (SEQ ID NO: 18).

[0053] A variant CasP6 protein has an amino acid sequence that is different by at least one amino acid (e.g., has a deletion, insertion, substitution, fusion) when compared to the amino acid sequence of the corresponding wild type CasP6 protein. A CasP6 protein that cleaves one strand but not the other of a double stranded target nucleic acid is referred to herein as a "nickase" (e.g., a "nickase CasP6"). A CasP6 protein that has substantially no nuclease activity is referred to herein as a dead CasP6 protein ("dCasP6") (with the caveat that nuclease activity can be provided by a heterologous polypeptide - a fusion partner - in the case of a chimeric CasP6 protein, which is described in more detail below). For any of the CasP6 variant proteins described herein (e.g., nickase CasP6, dCasP6, chimeric CasP6), the CasP6 variant can include a CasP6 protein sequence with the same parameters described above (e.g., domains that are present, percent identity, and the like).

[0054] In some cases, the CasP6 protein is a variant CasP6 protein, e.g., mutated relative to the naturally occurring catalytically active sequence, and exhibits reduced cleavage activity (e.g., exhibits 90%, or less, 80% or less, 70% or less, 60% or less, 50% or less, 40% or less, or 30% or less cleavage activity) when compared to the corresponding naturally occurring sequence. In some cases, such a variant CasP6 protein is a catalytically 'dead' protein (has substantially no cleavage activity) and can be referred to as a 'dCasP6.' In some cases, the variant CasP6 protein is a nickase (cleaves only one strand of a double stranded target nucleic acid, e.g., a double stranded target DNA). As described in more detail herein, in some cases, a CasP6 protein (in some case a CasP6 protein with wild type cleavage activity and in some cases a variant CasP6 with reduced cleavage activity, e.g., a dCasP6 or a nickase CasP6) is fused (conjugated) to a heterologous polypeptide that has an activity of interest (e.g., a catalytic activity of interest) to form a fusion protein (a chimeric CasP6 protein).

[0055] Conserved catalytic residues of CasP6A include D434 and E617 when numbered according to SEQ ID NO: 1. Conserved catalytic residues of CasP6B include D443 and E626 when numbered according to SEQ ID NO: 2. Conserved catalytic residues of CasP6C include D503 and E750 when numbered according to SEQ ID NO: 3. Corresponding conserved catalytic aspartate and glutamate residues of other CasP6 proteins can be identified by alignment of the CasP6 proteins set forth as SEQ ID NOs: 1-3. Mutations that can provide a dCasP6 protein include: (i) a substitution of the D434, D443, or D503 aspartate residues of CasP6A (SEQ ID NO:1), CasP6B (SEQ ID NO:2), or CasP6C (SEQ ID NO: 3), respectively, with another amino acid residue (*e.g.*, an alanine residue) as well as a similar substitutions of corresponding aspartate

residues in other CasP6 proteins; and/or (ii) a substitution of the E617, E626, or E750 glutamate residues of CasP6A (SEQ ID NO:1), CasP6B (SEQ ID NO:2), or CasP6C (SEQ ID NO: 3), respectively, with another residue (*e.g.,* an alanine residue) as well as a similar substitutions of corresponding glutamate residues in other CasP6 proteins.

[0056] Thus, in some cases, the CasP6 protein has reduced activity and one or more of the above described amino acids (or one or more corresponding amino acids of any CasP6 protein) are mutated (e.g., substituted with an alanine). In some cases, the variant CasP6 protein is a catalytically 'dead' protein (is catalytically inactive) and is referred to as 'dCasP6.' A dCasP6 protein can be fused to a fusion partner that provides an activity, and in some cases, the dCasP6 (e.g., one without a fusion partner that provides catalytic activity - but which can have an NLS when expressed in a eukaryotic cell) can bind to target DNA and can block RNA polymerase from translating from a target DNA or the function of other endogenous DNA binding or processing proteins. In some cases, the variant CasP6 protein is a nickase (cleaves only one strand of a double stranded target nucleic acid, e.g., a double stranded target DNA).

[0057] As noted above, in some cases, a CasP6 protein (in some cases a CasP6 protein with wild type cleavage activity and in some cases a variant CasP6 with reduced cleavage activity, e.g., a dCasP6 or a nickase CasP6) is fused (conjugated) to a heterologous polypeptide that has an activity of interest (e.g., a catalytic activity of interest) to form a fusion protein (a chimeric CasP6 protein). A heterologous polypeptide to which a CasP6 protein can be fused is referred to herein as a 'fusion partner.'

[0058] In some cases the fusion partner can modulate transcription (e.g., inhibit transcription, increase transcription) of a target DNA. For example, in some cases the fusion partner is a protein (or a domain from a protein) that inhibits transcription (e.g., a transcriptional repressor, a protein that functions via recruitment of transcription inhibitor proteins, modification of target DNA such as methylation, recruitment of a DNA modifier, modulation of histones associated with target DNA, recruitment of a histone modifier such as those that modify acetylation and/or methylation of histones, and the like). In some cases the fusion partner is a protein (or a domain from a protein) that increases transcription (e.g., a transcription activator, a protein that acts via recruitment of transcription activator proteins, modification of target DNA such as demethylation, recruitment of a DNA modifier, modulation of histones associated with target DNA, recruitment of a histone modifier such as those that modify acetylation and/or methylation of histones, and the like).

[0059] In some cases, a chimeric CasP6 protein includes a heterologous polypeptide that has enzymatic activity that modifies a target nucleic acid (e.g., nuclease activity, methyltransferase

activity, demethylase activity, DNA repair activity, DNA damage activity, deamination activity, dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity or glycosylase activity).

[0060] In some cases, a chimeric CasP6 protein includes a heterologous polypeptide that has enzymatic activity that modifies a polypeptide (e.g., a histone) associated with a target nucleic acid (e.g., methyltransferase activity, demethylase activity, acetyltransferase activity, deacetylase activity, kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity, deSUMOylating activity, ribosylation activity, deribosylation activity, myristoylation activity or demyristoylation activity).

[0061] Examples of proteins (or fragments thereof) that can be used in increase transcription include but are not limited to: transcriptional activators such as VP16, VP64, VP48, VP160, p65 subdomain (e.g., from NFkB), and activation domain of EDLL (SEQ ID NO: 19) and/or TAL activation domain (e.g., for activity in plants); histone lysine methyltransferases such as SET1A, SET1B, MLL1 to 5, ASH1, SYMD2, NSDl, and the like; histone lysine demethylases such as JHDM2a/b, UTX, JMJD3, and the like; histone acetyltransferases such as GCN5, PCAF, CBP, p300, TAF1, TIP60/PLIP, MOZ/MYST3, MORF/MYST4, SRCl, ACTR, P160, CLOCK, and the like; and DNA demethylases such as Ten-Eleven Translocation (TET) dioxygenase 1 (TET1CD), TET1, DME, DML1, DML2, ROS1, and the like.

[0062] Examples of proteins (or fragments thereof) that can be used in decrease transcription include but are not limited to: transcriptional repressors such as the Kriippel associated box (KRAB or SKD); KOX1 repression domain; the Mad mSIN3 interaction domain (SID); the ERF repressor domain (ERD), the SRDX repression domain (e.g., for repression in plants), and the like; histone lysine methyltransferases such as Pr-SET7/8, SUV4-20H1, RIZl, and the like; histone lysine demethylases such as JMJD2A/JHDM3A, JMJD2B, JMJD2C/GASC1, JMJD2D, JARID1A/RBP2, JARIDlB/PLU-1, JARID1C/SMCX, JARID1D/SMCY, and the like; histone lysine deacetylases such as HDAC1, HDAC2, HDAC3, HDAC8, HDAC4, HDAC5, HDAC7, HDAC9, SIRT1, SIRT2, HDAC11, and the like; DNA methylases such as Hhal DNA m5c-methyltransferase (M.Hhal), DNA methyltransferase 1 (DNMT1), DNA methyltransferase 3a (DNMT3a), DNA methyltransferase 3b (DNMT3b), METI, DRM3 (plants), ZMET2, CMT1, CMT2 (plants), and the like; and periphery recruitment elements such as Lamin A, Lamin B, and the like.

[0063] In some cases the fusion partner has enzymatic activity that modifies the target nucleic acid (e.g., ssRNA, dsRNA, ssDNA, dsDNA). Examples of enzymatic activity that can be

provided by the fusion partner include but are not limited to: nuclease activity such as that provided by a restriction enzyme (e.g., Fokl nuclease), methyltransferase activity such as that provided by a methyltransferase (e.g., Hhal DNA m5c-methyltransferase, M.Hhal), DNA methyltransferase 1 (DNMT1), DNA methyltransferase 3a (DNMT3a), DNA methyltransferase 3b (DNMT3b), MET1, DRM3 (plants), ZMET2, CMT1, CMT2 (plants), and the like); demethylase activity such as that provided by a demethylase (e.g., Ten-Eleven Translocation (TET) dioxygenase 1 (TET1CD), TET1, DME, DML1, DML2, ROS 1, and the like), DNA repair activity, DNA damage activity, deamination activity such as that provided by a deaminase (e.g., a cytosine deaminase enzyme such as rat APOBEC1), dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity such as that provided by an integrase and/or resolvase (e.g., Gin invertase such as the hyperactive mutant of the Gin invertase, GinH106Y; human immunodeficiency virus type 1 integrase (IN); Tn3 resolvase; and the like), transposase activity, recombinase activity such as that provided by a recombinase (e.g., catalytic domain of Gin recombinase), polymerase activity, ligase activity, helicase activity, photolyase activity, and glycosylase activity).

[0064] In some cases the fusion partner has enzymatic activity that modifies a protein associated with the target nucleic acid (e.g., ssRNA, dsRNA, ssDNA, dsDNA) (e.g., a histone, an RNA binding protein, a DNA binding protein, and the like). Examples of enzymatic activity (that modifies a protein associated with a target nucleic acid) that can be provided by the fusion partner include but are not limited to: methyltransferase activity such as that provided by a histone methyltransferase (HMT) (e.g., suppressor of variegation 3-9 homolog 1 (SUV39H1, also known as KMTIA), euchromatic histone lysine methyltransferase 2 (G9A, also known as KMT1C and EHMT2), SUV39H2, ESET/SETDB 1, and the like, SET1A, SET1B, MLL1 to 5, ASH1, SYMD2, NSD1, DOT1L, Pr-SET7/8, SUV4-20H1, EZH2, RIZl), demethylase activity such as that provided by a histone demethylase (e.g., Lysine Demethylase 1A (KDM1A also known as LSD1), JHDM2a/b, JMJD2A/JHDM3A, JMJD2B, JMJD2C/GASC1, JMJD2D, JARID1A/RBP2, JARIDlB/PLU-1, JARID1C/SMCX, JARID1D/SMCY, UTX, JMJD3, and the like), acetyltransferase activity such as that provided by a histone acetylase transferase (e.g., catalytic core/fragment of the human acetyltransferase p300, GCN5, PCAF, CBP, TAF1, TIP60/PLIP, MOZ/MYST3, MORF/MYST4, HB01/MYST2, HMOF/MYST1, SRC1, ACTR, P160, CLOCK, and the like), deacetylase activity such as that provided by a histone deacetylase (e.g., HDAC1, HDAC2, HDAC3, HDAC8, HDAC4, HDAC5, HDAC7, HDAC9, SIRT1, SIRT2, HDAC11, and the like), kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity,

deSUMOylating activity, ribosylation activity, deribosylation activity, myristoylation activity, and demyristoylation activity.

[0065] An additional examples of a suitable fusion partners are dihydrofolate reductase (DHFR) destabilization domain (e.g., to generate a chemically controllable chimeric CasP6 protein), and a chloroplast transit peptide. Suitable chloroplast transit peptides include, but are not limited to: SEQ ID NO: 20; SEQ ID NO: 21; SEQ ID NO: 22; SEQ ID NO: 23; SEQ ID NO: 24; SEQ ID NO: 25; SEQ ID NO: 26; SEQ ID NO: 27; SEQ ID NO: 28; SEQ ID NO: 29; and SEQ ID NO: 30.

[0066] In some case, a CasP6 fusion polypeptide of the present disclosure comprises: a) a CasP6 polypeptide of the present disclosure; and b) a chloroplast transit peptide. Thus, for example, a CRISPR- CasP6 complex can be targeted to the chloroplast. In some cases, this targeting may be achieved by the presence of an N-terminal extension, called a chloroplast transit peptide (CTP) or plastid transit peptide. Chromosomal transgenes from bacterial sources must have a sequence encoding a CTP sequence fused to a sequence encoding an expressed polypeptide if the expressed polypeptide is to be compartmentalized in the plant plastid (e.g. chloroplast). Accordingly, localization of an exogenous polypeptide to a chloroplast is often 1 accomplished by means of operably linking a polynucleotide sequence encoding a CTP sequence to the 5' region of a polynucleotide encoding the exogenous polypeptide. The CTP is removed in a processing step during translocation into the plastid. Processing efficiency may, however, be affected by the amino acid sequence of the CTP and nearby sequences at the NH 2 terminus of the peptide. Other options for targeting to the chloroplast which have been described are the maize cab-m7 signal sequence (U.S. Pat. No. 7,022,896, WO 97/41228) a pea glutathione reductase signal sequence (WO 97/41228) and the CTP described in US2009029861.

[0067] The CasP6 polypeptide disclosed herein can further comprise at least one plastid targeting signal peptide, at least one mitochondrial targeting signal peptide, or a signal peptide targeting the CasP6 polypeptide to both plastids and mitochondria. Plastid, mitochondrial, and dual-targeting signal peptide localization signals are known in the art (see, e.g., Nassoury and Morse (2005) Biochim Biophys Acta 1743:5-19; Kunze and Berger (2015) Front Physiol dx.doi.org/10.3389/fphys.2015.00259; Herrmann and Neupert (2003) IUBMB Life 55:219-225; Soll (2002) Curr Opin Plant Biol 5:529-535; Carrie and Small (2013) Biochim Biophys Acta 1833:253-259; Carrie et al. (2009) FEBS J 276:1187-1195; Silva-Filho (2003) Curr Opin Plant Biol 6:589-595; Peeters and Small (2001) Biochim Biophys Acta 1541:54-63; Murcha et al. (2014) J Exp Bot 65:6301-6335; Mackenzie (2005) Trends Cell Biol 15:548-554; Glaser et al. (1998) Plant Mol Biol 38:311-338). The plastid, mitochondrial, or dual-targeting signal peptide

can be located at the N-terminus, the C-terminus, or in an internal location of the CasP6 polypeptide.

[0068] In some cases, a CasP6 fusion polypeptide of the present disclosure can comprise: a) a CasP6 polypeptide of the present disclosure; and b) an endosomal escape peptide. In some cases, an endosomal escape polypeptide comprises the amino acid sequence of SEQ ID NO: 31, wherein each X is independently selected from lysine, histidine, and arginine. In some cases, an endosomal escape polypeptide comprises the amino acid sequence of SEQ ID NO: 32.

[0069] For examples of some of the above fusion partners (and more) used in the context of fusions with Cas9, Zinc Finger, and/or TALE proteins (for site specific target nucleic modification, modulation of transcription, and/or target protein modification, e.g., histone modification), see, e.g.: Nomura et al J Am Chem Soc. 2007 Jul 18;129(28):8676-7; Rivenbark et al., Epigenetics. 2012 Apr;7(4):350-60; Nucleic Acids Res. 2016 Jul 8;44(12):5615-28; Gilbert et. al, Cell. 2013 Jul 18;154(2):442-51 ; Kearns et al, Nat Methods. 2015 May;12(5):401-3; Mendenhall et. al, Nat Biotechnol. 2013 Dec;31(12): 1133-6; Hilton et. al., Nat Biotechnol. 2015 May;33(5):510-7; Gordley et. al., Proc Natl Acad Sci U S A. 2009 Mar 31 ;106(13):5053-8; Akopian et. al., Proc Natl Acad Sci U S A. 2003 Jul 22;100(15):8688-91 ; Tan et., al., J Virol. 2006 Feb; 80(4): 1939-48; Tan et. al., Proc Natl Acad Sci U S A. 2003 Oct 14; 100(21): 11997-2002; Papworth et. al., Proc Natl Acad Sci U S A. 2003 Feb 18;100(4): 1621-6; Sanjana et. al., Nat Protoc. 2012 Jan 5;7(l): 171-92; Beerli et. al., Proc Natl Acad Sci U S A. 1998 Dec 8;95(25): 14628-33; Snowden et. al., Curr Biol. 2002 Dec 23;12(24):2159-66; Xu et.al., Xu et. al., Cell Discov. 2016 May 3;2: 16009; Komor et al., Nature. 2016 Apr 20;533(7603):420-4; Chaikind et. al., Nucleic Acids Res. 2016 Aug 11 ; Choudhury at. al., Oncotarget. 2016 Jun 23; Du et. al., Cold Spring Harb Protoc. 2016 Jan 4; Pham et. al, Methods Mol Biol. 2016;1358:43-57; Balboa et al., Stem Cell Reports. 2015 Sep 8;5(3):448-59; Hara et. al., Sci Rep. 2015 Jun 9;5: 11221; Piatek et. al., Plant Biotechnol J. 2015 May;13(4):578-89; Hu et al., Nucleic Acids Res. 2014 Apr;42(7):4375-90; Cheng et. al., Cell Res. 2013 Oct;23(10): 1163-71 ; Cheng et. al, Cell Res. 2013 Oct;23(10):l 163-71; and Maeder et. al., Nat Methods. 2013 Oct;10(10):977-9.

[0070] Additional suitable heterologous polypeptide include, but are not limited to, a polypeptide that directly and/or indirectly provides for increased transcription and/or translation of a target nucleic acid (e.g., a transcription activator or a fragment thereof, a protein or fragment thereof that recruits a transcription activator, a small molecule/drug-responsive transcription and/or translation regulator, a translation-regulating protein, etc.). Non-limiting examples of heterologous polypeptides to accomplish increased or decreased transcription include transcription activator and transcription repressor domains. In some such cases, a chimeric CasP6

polypeptide is targeted by the guide nucleic acid (guide RNA) to a specific location (i.e., sequence) in the target nucleic acid and exerts locus-specific regulation such as blocking RNA polymerase binding to a promoter (which selectively inhibits transcription activator function), and/or modifying the local chromatin status (e.g., when a fusion sequence is used that modifies the target nucleic acid or modifies a polypeptide associated with the target nucleic acid). In some cases, the changes are transient (e.g., transcription repression or activation). In some cases, the changes are inheritable (e.g., when epigenetic modifications are made to the target nucleic acid or to proteins associated with the target nucleic acid, e.g., nucleosomal histones).

[0071] Non-limiting examples of heterologous polypeptides for use when targeting ssRNA target nucleic acids include but are not limited to: splicing factors (e.g., RS domains); protein translation components (e.g., translation initiation, elongation, and/or release factors; e.g., eIF4G); RNA methylases; RNA editing enzymes (e.g., RNA deaminases, e.g., adenosine deaminase acting on RNA (ADAR), including A to I and/or C to U editing enzymes); helicases; RNA-binding proteins; and the like. It is understood that a heterologous polypeptide can include the entire protein or in some cases can include a fragment of the protein (e.g., a functional domain).

[0072] The heterologous polypeptide of a subject chimeric CasP6 polypeptide can be any domain capable of interacting with ssRNA (which, for the purposes of this disclosure, includes intramolecular and/or intermolecular secondary structures, e.g., double-stranded RNA duplexes such as hairpins, stem- loops, etc.), whether transiently or irreversibly, directly or indirectly, including but not limited to an effector domain selected from the group comprising: Endonucleases (for example RNase III, the CRR22 DYW domain, Dicer, and PIN (PilT N-terminus) domains from proteins such as SMG5 and SMG6); proteins and protein domains responsible for stimulating RNA cleavage (for example CPSF, CstF, CFIm and CFIIm); Exonucleases (for example XRN-1 or Exonuclease T); Deadenylases (for example HNT3); proteins and protein domains responsible for nonsense mediated RNA decay (for example UPF1, UPF2, UPF3, UPF3b, RNP SI, Y14, DEK, REF2, and SRml60); proteins and protein domains responsible for stabilizing RNA (for example PABP) ; proteins and protein domains responsible for repressing translation (for example Ago2 and Ago4); proteins and protein domains responsible for stimulating translation (for example Staufen); proteins and protein domains responsible for (e.g., capable of) modulating translation (e.g., translation factors such as initiation factors, elongation factors, release factors, etc., e.g., eIF4G); proteins and protein domains responsible for polyadenylation of RNA (for example PAP1, GLD-2, and Star- PAP) ; proteins and protein domains responsible for polyuridinylation of RNA (for example CI Dl and terminal

uridylate transferase) ; proteins and protein domains responsible for RNA localization (for example from IMPl, ZBPl, She2p, She3p, and Bicaudal-D); proteins and protein domains responsible for nuclear retention of RNA (for example Rrp6); proteins and protein domains responsible for nuclear export of RNA (for example TAP, NXF1, THO, TREX, REF, and Aly) ; proteins and protein domains responsible for repression of RNA splicing (for example PTB, Sam68, and hnRNP Al) ; proteins and protein domains responsible for stimulation of RNA splicing (for example Serine/ Arginine-rich (SR) domains) ; proteins and protein domains responsible for reducing the efficiency of transcription (for example FUS (TLS)); and proteins and protein domains responsible for stimulating transcription (for example CDK7 and HIV Tat). Alternatively, the effector domain may be selected from the group comprising Endonucleases; proteins and protein domains capable of stimulating RNA cleavage; Exonucleases; Deadenylases; proteins and protein domains having nonsense mediated RNA decay activity; proteins and protein domains capable of stabilizing RNA; proteins and protein domains capable of repressing translation; proteins and protein domains capable of stimulating translation; proteins and protein domains capable of modulating translation (e.g., translation factors such as initiation factors, elongation factors, release factors, etc., e.g., eIF4G); proteins and protein domains capable of polyadenylation of RNA; proteins and protein domains capable of polyuridinylation of RNA; proteins and protein domains having RNA localization activity; proteins and protein domains capable of nuclear retention of RNA; proteins and protein domains having RNA nuclear export activity; proteins and protein domains capable of repression of RNA splicing; proteins and protein domains capable of stimulation of RNA splicing; proteins and protein domains capable of reducing the efficiency of transcription ; and proteins and protein domains capable of stimulating transcription. Another suitable heterologous polypeptide is a PUF RNA-binding domain, which is described in more detail in WO2012068627, which is hereby incorporated by reference in its entirety.

[0073] Some RNA splicing factors that can be used (in whole or as fragments thereof) as heterologous polypeptides for a chimeric CasP6 polypeptide have modular organization, with separate sequence-specific RNA binding modules and splicing effector domains. For example, members of the Serine/ Arginine-rich (SR) protein family contain N-terminal RNA recognition motifs (RRMs) that bind to exonic splicing enhancers (ESEs) in pre-mRNAs and C-terminal RS domains that promote exon inclusion. As another example, the hnRNP protein hnRNP Al binds to exonic splicing silencers (ESSs) through its RRM domains and inhibits exon inclusion through a C-terminal Glycine -rich domain. Some splicing factors can regulate alternative use of splice site (ss) by binding to regulatory sequences between the two alternative sites. For example,

ASF/SF2 can recognize ESEs and promote the use of intron proximal sites, whereas hnRNP Al can bind to ESSs and shift splicing towards the use of intron distal sites. One application for such factors is to generate ESFs that modulate alternative splicing of endogenous genes, particularly disease associated genes. For example, Bcl-x pre-mRNA produces two splicing isoforms with two alternative 5' splice sites to encode proteins of opposite functions. The long splicing isoform Bcl-xL is a potent apoptosis inhibitor expressed in long-lived postmitotic cells and is up-regulated in many cancer cells, protecting cells against apoptotic signals. The short isoform Bcl-xS is a pro-apoptotic isoform and expressed at high levels in cells with a high turnover rate (e.g., developing lymphocytes). The ratio of the two Bcl-x splicing isoforms is regulated by multiple cc -elements that are located in either the core exon region or the exon extension region (i.e., between the two alternative 5' splice sites). For more examples, see WO2010075303, which is hereby incorporated by reference in its entirety.

[0074] Further suitable fusion partners include, but are not limited to proteins (or fragments thereof) that are boundary elements (e.g., CTCF), proteins and fragments thereof that provide periphery recruitment (e.g., Lamin A, Lamin B, etc.), protein docking elements (e.g., FKBP/FRB, Pill/Abyl, etc.).

[0075] Examples of various additional suitable heterologous polypeptide (or fragments thereof) for a subject chimeric CasP6 polypeptide include, but are not limited to those described in the following applications (which publications are related to other CRISPR endonucleases such as Cas9, but the described fusion partners can also be used with CasP6 instead): PCT patent applications: WO2010075303, WO2012068627, and WO2013155555, and can be found, for example, in U.S. patents and patent applications: 8,906,616; 8,895,308; 8,889,418; 8,889,356; 8,871,445; 8,865,406; 8,795,965; 8,771,945; 8,697,359; 20140068797; 20140170753; 20140179006; 20140179770; 20140186843; 20140186919; 20140186958; 20140189896; 20140227787; 20140234972; 20140242664; 20140242699; 20140242700; 20140242702; 20140248702; 20140256046; 20140273037; 20140273226; 20140273230; 20140273231 ; 20140273232; 20140273233; 20140273234; 20140273235; 20140287938; 20140295556; 20140295557; 20140298547; 20140304853; 20140309487; 20140310828; 20140310830; 20140315985; 20140335063; 20140335620; 20140342456; 20140342457; 20140342458; 20140349400; 20140349405; 20140356867; 20140356956; 20140356958; 20140356959; 20140357523; 20140357530; 20140364333; and 20140377868; all of which are hereby incorporated by reference in their entirety.

[0076] In some cases, a heterologous polypeptide (a fusion partner) provides for subcellular localization, i.e., the heterologous polypeptide contains a subcellular localization sequence (e.g.,

a nuclear localization signal (NLS) for targeting to the nucleus, a sequence to keep the fusion protein out of the nucleus, e.g., a nuclear export sequence (NES), a sequence to keep the fusion protein retained in the cytoplasm, a mitochondrial localization signal for targeting to the mitochondria, a chloroplast localization signal for targeting to a chloroplast, an ER retention signal, and the like). In some embodiments, a CasP6 fusion polypeptide does not include a NLS so that the protein is not targeted to the nucleus (which can be advantageous, e.g., when the target nucleic acid is an RNA that is present in the cytosol). In some embodiments, the heterologous polypeptide can provide a tag (i.e., the heterologous polypeptide is a detectable label) for ease of tracking and/or purification (e.g., a fluorescent protein, e.g., green fluorescent protein (GFP), YFP, RFP, CFP, mCherry, tdTomato, and the like; a histidine tag, e.g., a 6XHis tag; a hemagglutinin (HA) tag; a FLAG tag; a Myc tag; and the like).

[0077] In some cases a CasP6 protein (e.g., a wild type CasP6 protein, a variant CasP6 protein, a chimeric CasP6 protein, a dCasP6 protein, a chimeric CasP6 protein where the CasP6 portion has reduced nuclease activity - such as a dCasP6 protein fused to a fusion partner, and the like) includes (is fused to) a nuclear localization signal (NLS) (e.g, in some cases 2 or more, 3 or more, 4 or more, or 5 or more NLSs). Thus, in some cases, a CasP6 polypeptide includes one or more NLSs (e.g., 2 or more, 3 or more, 4 or more, or 5 or more NLSs). In some cases, one or more NLSs (2 or more, 3 or more, 4 or more, or 5 or more NLSs) are positioned at or near (e.g., within 50 amino acids of) the N-terminus and/or the C- terminus. In some cases, one or more NLSs (2 or more, 3 or more, 4 or more, or 5 or more NLSs) are positioned at or near (e.g., within 50 amino acids of) the N-terminus. In some cases, one or more NLSs (2 or more, 3 or more, 4 or more, or 5 or more NLSs) are positioned at or near (e.g., within 50 amino acids of) the C-terminus. In some cases, one or more NLSs (3 or more, 4 or more, or 5 or more NLSs) are positioned at or near (e.g., within 50 amino acids of) both the N-terminus and the C-terminus. In some cases, an NLS is positioned at the N-terminus and an NLS is positioned at the C-terminus.

[0078] In some cases a CasP6 protein (e.g., a wild type CasP6 protein, a variant CasP6 protein, a chimeric CasP6 protein, a dCasP6 protein, a chimeric CasP6 protein where the CasP6 portion has reduced nuclease activity - such as a dCasP6 protein fused to a fusion partner, and the like) includes (is fused to) between 1 and 10 NLSs (e.g., 1-9, 1-8, 1-7, 1-6, 1-5, 2-10, 2-9, 2-8, 2-7, 2-6, or 2-5 NLSs). In some cases a CasP6 protein (e.g., a wild type CasP6 protein, a variant CasP6 protein, a chimeric CasP6 protein, a dCasP6 protein, a chimeric CasP6 protein where the CasP6 portion has reduced nuclease activity - such as a dCasP6 protein fused to a fusion partner, and the like) includes (is fused to) between 2 and 5 NLSs (e.g., 2-4, or 2- 3 NLSs).

[0079] Non-limiting examples of NLSs include an NLS sequence derived from: the NLS of the

SV40 virus large T-antigen, having the amino acid sequence of SEQ ID NO: 33; the NLS from nucleoplasmin (e.g., the nucleoplasmin bipartite NLS with the sequence of SEQ ID NO: 34); the c-myc NLS having the amino acid sequence of SEQ ID NO: 35 or of SEQ ID NO: 36; the hRNPAl M9 NLS having the sequence of SEQ ID NO: 37; the sequence of SEQ ID NO: 38 of the IBB domain from importin-alpha; the sequences of SEQ ID NO: 39 and of SEQ ID NO: 40 of the myoma T protein; the sequence of SEQ ID NO: 41 of human p53; the sequence of SEQ ID NO: 42 of mouse c-abl IV; the sequences of SEQ ID NO: 43 and of SEQ ID NO: 44 of the influenza virus NS1; the sequence of SEQ ID NO: 45 of the Hepatitis virus delta antigen; the sequence of SEQ ID NO: 46 of the mouse Mxl protein; the sequence of SEQ ID NO: 47 of the human poly(ADP-ribose) polymerase; and the sequence of SEQ ID NO: 48 of the steroid hormone receptors (human) glucocorticoid. In general, NLS (or multiple NLSs) are of sufficient strength to drive accumulation of the CasP6 protein in a detectable amount in the nucleus of a eukaryotic cell. Detection of accumulation in the nucleus may be performed by any suitable technique. For example, a detectable marker may be fused to the CasP6 protein such that location within a cell may be visualized. Cell nuclei may also be isolated from cells, the contents of which may then be analyzed by any suitable process for detecting protein, such as immunohistochemistry, Western blot, or enzyme activity assay. Accumulation in the nucleus may also be determined indirectly.

[0080] In some cases, a CasP6 fusion polypeptide includes a "Protein Transduction Domain" or PTD (also known as a CPP - cell penetrating peptide), which refers to a polypeptide, polynucleotide, carbohydrate, or organic or inorganic compound that facilitates traversing a lipid bilayer, micelle, cell membrane, organelle membrane, or vesicle membrane. A PTD attached to another molecule, which can range from a small polar molecule to a large macromolecule and/or a nanoparticle, facilitates the molecule traversing a membrane, for example going from extracellular space to intracellular space, or cytosol to within an organelle. In some embodiments, a PTD is covalently linked to the amino terminus a polypeptide (e.g., linked to a wild type CasP6 to generate a fusino protein, or linked to a variant CasP6 protein such as a dCasP6, nickase CasP6, or chimeric CasP6 protein to generate a fusion protein). In some embodiments, a PTD is covalently linked to the carboxyl terminus of a polypeptide (e.g., linked to a wild type CasP6 to generate a fusino protein, or linked to a variant CasP6 protein such as a dCasP6, nickase CasP6, or chimeric CasP6 protein to generate a fusion protein). In some cases, the PTD is inserted internally in the CasP6 fusion polypeptide (i.e., is not at the N- or C-terminus of the CasP6 fusion polypeptide) at a suitable insertion site. In some cases, a subject CasP6 fusion polypeptide includes (is conjugated to, is fused to) one or more PTDs (e.g., two or more, three or more, four

or more PTDs). In some cases a PTD includes a nuclear localization signal (NLS) (e.g, in some cases 2 or more, 3 or more, 4 or more, or 5 or more NLSs). Thus, in some cases, a CasP6 fusion polypeptide includes one or more NLSs (e.g., 2 or more, 3 or more, 4 or more, or 5 or more NLSs). In some embodiments, a PTD is covalently linked to a nucleic acid (e.g., a CasP6 guide nucleic acid, a polynucleotide encoding a CasP6 guide nucleic acid, a polynucleotide encoding a CasP6 fusion polypeptide, a donor polynucleotide, etc.). Examples of PTDs include but are not limited to a minimal undecapeptide protein transduction domain (corresponding to residues 47-57 of HIV-1 TAT comprising SEQ ID NO: 49); a polyarginine sequence comprising a number of arginines sufficient to direct entry into a cell (e.g., 3, 4, 5, 6, 7, 8, 9, 10, or 10-50 arginines); a VP22 domain (Zender et al. (2002) Cancer Gene Ther. 9(6):489-96); an Drosophila Antennapedia protein transduction domain (Noguchi et al. (2003) Diabetes 52(7): 1732-1737); a truncated human calcitonin peptide (Trehin et al. (2004) Pharm. Research 21: 1248-1256); polylysine (Wender et al. (2000) Proc. Natl. Acad. Sci. USA 97: 13003-13008); SEQ ID NO: 50; Transportan SEQ ID NO: 51; SEQ ID NO: 52; and SEQ ID NO: 53. Exemplary PTDs include but are not limited to, , SEQ ID NO: 54; an arginine homopolymer of from 3 arginine residues to 50 arginine residues. Exemplary PTD domain amino acid sequences include, but are not limited to, any of the following:; SEQ ID NO: 55; SEQ ID NO: 56; SEQ ID NO: 57; and SEQ ID NO: 58. In some embodiments, the PTD is an activatable CPP (ACPP) (Aguilera et al. (2009) Integr Biol (Camb) June; 1(5-6): 371-381). ACPPs comprise a polycationic CPP (e.g., Arg9 or "R9") connected via a cleavable linker to a matching polyanion (e.g., Glu9 or "E9"), which reduces the net charge to nearly zero and thereby inhibits adhesion and uptake into cells. Upon cleavage of the linker, the polyanion is released, locally unmasking the polyarginine and its inherent adhesiveness, thus "activating" the ACPP to traverse the membrane.

[0081] In some embodiments, a subject CasP6 protein can fused to a fusion partner via a linker polypeptide (e.g., one or more linker polypeptides). The linker polypeptide may have any of a variety of amino acid sequences. Proteins can be joined by a spacer peptide, generally of a flexible nature, although other chemical linkages are not excluded. Suitable linkers include polypeptides of between 4 amino acids and 40 amino acids in length, or between 4 amino acids and 25 amino acids in length. These linkers can be produced by using synthetic, linker-encoding oligonucleotides to couple the proteins, or can be encoded by a nucleic acid sequence encoding the fusion protein. Peptide linkers with a degree of flexibility can be used. The linking peptides may have virtually any amino acid sequence, bearing in mind that the preferred linkers will have a sequence that results in a generally flexible peptide. The use of small amino acids, such as glycine and alanine, are of use in creating a flexible peptide. The creation of such sequences is

routine to those of skill in the art. A variety of different linkers are commercially available and are considered suitable for use.

[0082] Examples of linker polypeptides include glycine polymers $(G)_n$, glycine-serine polymers (including, for example, $(GS)_n$, $GSGGS_n$ (SEQ ID NO: 59), $GGSGGS_n$ (SEQ ID NO: 60), and $GGGS_n$ (SEQ ID NO: 61), where n is an integer of at least one), glycine-alanine polymers, alanine-serine polymers. Exemplary linkers can comprise amino acid sequences including, but not limited to, GGSG (SEQ ID NO: 62), GGSGG (SEQ ID NO: 63), GSGSG (SEQ ID NO: 64), GSGGG (SEQ ID NO: 65), GGGSG (SEQ ID NO: 66), GSSSG (SEQ ID NO: 67), and the like. The ordinarily skilled artisan will recognize that design of a peptide conjugated to any desired element can include linkers that are all or partially flexible, such that the linker can include a flexible linker as well as one or more portions that confer less flexible structure.

[0083] In some cases, a CasP6 polypeptide of the present disclosure comprises a detectable label. Suitable detectable labels and/or moieties that can provide a detectable signal can include, but are not limited to, an enzyme, a radioisotope, a member of a specific binding pair; a fluorophore; a fluorescent protein; a quantum dot; and the like.

[0084] Suitable fluorescent proteins include, but are not limited to, green fluorescent protein (GFP) or variants thereof, blue fluorescent variant of GFP (BFP), cyan fluorescent variant of GFP (CFP), yellow fluorescent variant of GFP (YFP), enhanced GFP (EGFP), enhanced CFP (ECFP), enhanced YFP (EYFP), GFPS65T, Emerald, Topaz (TYFP), Venus, Citrine, mCitrine, GFPuv, destabilised EGFP (dEGFP), destabilised ECFP (dECFP), destabilised EYFP (dEYFP), mCFPm, Cerulean, T-Sapphire, CyPet, YPet, mKO, HcRed, t-HcRed, DsRed, DsRed2, DsRed-monomer, J-Red, dimer2, t-dimer2(12), mRFPl, pocilloporin, Renilla GFP, Monster GFP, paGFP, Kaede protein and kindling protein, Phycobiliproteins and Phycobiliprotein conjugates including B-Phycoerythrin, R-Phycoerythrin and Allophycocyanin. Other examples of fluorescent proteins include mHoneydew, mBanana, mOrange, dTomato, tdTomato, mTangerine, mStrawberry, mCherry, mGrapel, mRaspberry, mGrape2, mPlum (Shaner et al. (2005) Nat. Methods 2:905-909), and the like. Any of a variety of fluorescent and colored proteins from Anthozoan species, as described in, e.g., Matz et al. (1999) Nature Biotechnol. 17:969-973, is suitable for use.

[0085] Suitable enzymes include, but are not limited to, horse radish peroxidase (HRP), alkaline phosphatase (AP), beta-galactosidase (GAL), glucose-6-phosphate dehydrogenase, beta-N-acetylglucosaminidase, β-glucuronidase, invertase, Xanthine Oxidase, firefly luciferase, glucose oxidase (GO), and the like.

[0086] A CasP6 protein binds to target nucleic acid at a target sequence defined by the region of complementarity between the nucleic acid-targeting RNA and the target nucleic acid. As is the

case for many CRISPR endonucleases, site-specific binding (and/or cleavage) of a double stranded target DNA occurs at locations determined by (i) base-pairing complementarity between the guide RNA and the target nucleic acid; and sometimes (ii) a short motif referred to as the protospacer adjacent motif (PAM) in the target DNA or as the Protospacer Flanking Sequence (PFS) in the target RNA.

[0087] In some embodiments, a PAM for a CasP6 protein is immediately 5' of the target sequence of the non-complementary strand of the target DNA (the complementary strand hybridizes to the guide sequence of the guide RNA while the non-complementary strand does not directly hybridize with the guide RNA and is the reverse complement of the complementary strand). In some embodiments (e.g., when CasP6 as described herein is used), a PAM sequence of the non- complementary strand is T-rich. Examples of PAM sequences include, but are not limited to, TTT, TTN, CTN, TCN, CCN, TTTN, TCTN, TTCN, CTTN, ATTN, TCCN, TTGN, GTTN, CCCN, CCTN, TTAN, TCGN, CTCN, ACTN, GCTN, TCAN, GCCN, and CCGN (wherein N is defined as any nucleotide). When the target is RNA, the adjacent motif is sometimes called Protospacer Flanking Sequence (PFS), and it can be shorter, such as one or two residues in length.

[0088] In some cases, different CasP6 proteins (i.e., CasP6 proteins from various species) may be advantageous to use in the various provided methods in order to capitalize on various enzymatic characteristics of the different CasP6 proteins (e.g., for different PAM sequence preferences; for increased or decreased enzymatic activity; for an increased or decreased level of cellular toxicity; to change the balance between NHEJ, homology-directed repair, single strand breaks, double strand breaks, etc.; to take advantage of a short total sequence; and the like). CasP6 proteins from different species may require different PAM sequences in the target DNA or RNA. Thus, for a particular CasP6 protein of choice, the PAM sequence requirement may be different than the sequence described above. Various methods (including in silico and/or wet lab methods) for identification of the appropriate PAM sequence are known in the art and are routine, and any convenient method can be used. A PAM sequence can be identified using a PAM depletion assay.

[0089] A nucleic acid molecule that binds to a CasP6 protein, forming a ribonucleoprotein complex (RNP), and targets the complex to a specific location within a target nucleic acid (e.g., a target DNA or a target RNA) is referred to herein as a "CasP6 guide RNA" or simply as a "guide RNA." It is to be understood that in some cases, a hybrid DNA/RNA can be made such that a CasP6 guide RNA includes DNA bases in addition to RNA bases, but the term "CasP6 guide RNA" is still used to encompass such a molecule herein.

[0090] A CasP6 guide RNA can be said to include two segments, a targeting segment and a protein-binding segment. The targeting segment of a CasP6 guide RNA includes a nucleotide sequence (a guide sequence) that is complementary to (and therefore hybridizes with) a specific sequence (a target site) within a target nucleic acid (e.g., a target ssRNA, a target ssDNA, the complementary strand of a double stranded target DNA, etc.). Site-specific binding and/or cleavage of a target nucleic acid (e.g., genomic DNA or a target RNA) can occur at locations (e.g., target sequence of a target locus) determined by base-pairing complementarity between the CasP6 guide RNA (the guide sequence of the CasP6 guide RNA) and the target nucleic acid.

[0091] The protein-binding segment (or "protein-binding sequence") interacts with (binds to) a CasP6 polypeptide.

[0092] In some cases the protein-binding segment is made up of a short sequence of 17-23 nucleotides, such as a sequence of 21 or 22 nucleotides. This protein binding segment forms a double-stranded RNA duplex of five paired residues in length. The 5' terminus has about three residues upstream from the first RNA duplexed residue. A stem structure of 4-5 residues separates the double stranded regions.

[0093] In some cases the protein-binding segment of a subject CasP6 guide RNA includes two complementary stretches of nucleotides that hybridize to one another to form a double stranded RNA duplex (dsRNA duplex).

[0094] A CasP6 guide RNA and a CasP6 protein, e.g., a fusion CasP6 polypeptide, form a complex (e.g., bind via non-covalent interactions). The CasP6 guide RNA provides target specificity to the complex by including a targeting segment, which includes a guide sequence (a nucleotide sequence that is complementary to a sequence of a target nucleic acid). The CasP6 protein of the complex provides the site- specific activity (e.g., cleavage activity provided by the CasP6 protein and/or an activity provided by the fusion partner in the case of a chimeric CasP6 protein). In other words, the CasP6 protein is guided to a target nucleic acid sequence (e.g. a target sequence) by virtue of its association with the CasP6 guide RNA.

[0095] The "guide sequence" also referred to as the "targeting sequence" of a CasP6 guide RNA can be made so that the CasP6 guide RNA can target a CasP6 protein (e.g., a naturally occurring CasP6 protein, a fusion CasP6 polypeptide (chimeric CasP6), and the like) to any desired sequence of any desired target nucleic acid, with the exception (e.g., as described herein) that the PAM sequence can be taken into account. Thus, for example, a CasP6 guide RNA can have a guide sequence with complementarity to (e.g., can hybridize to) a sequence in a nucleic acid in a eukaryotic cell, e.g., a viral nucleic acid, a eukaryotic nucleic acid (e.g., a eukaryotic chromosome, chromosomal sequence, a eukaryotic RNA, etc.), and the like.

[0096] A subject CasP6 guide RNA can also be said to include an "activator" and a "targeter"(e.g., an "activator-RNA" and a "targeter-RNA," respectively). When the "activator" and a "targeter" are two separate molecules the guide RNA is referred to herein as a "dual guide RNA", a "dgRNA," a "double-molecule guide RNA", or a "two-molecule guide RNA." (e.g., a "CasP6 dual guide RNA"). In some embodiments, the activator and targeter are covalently linked to one another (e.g., via intervening nucleotides) and the guide RNA is referred to herein as a "single guide RNA", an "sgRNA," a "single- molecule guide RNA," or a "one-molecule guide RNA" (e.g., a "CasP6 single guide RNA"). Thus, a subject CasP6 single guide RNA comprises a targeter (e.g., targeter-RNA) and an activator (e.g., activator- RNA) that are linked to one another (e.g., by intervening nucleotides), and hybridize to one another to form the double stranded RNA duplex (dsRNA duplex) of the protein-binding segment of the guide RNA, thus resulting in a stem-loop structure. Thus, the targeter and the activator each have a duplex-forming segment, where the duplex forming segment of the targeter and the duplex-forming segment of the activator have complementarity with one another and hybridize to one another.

[0097] In some embodiments, the linker of a CasP6 single guide RNA is a stretch of nucleotides. In some cases, the targeter and activator of a CasP6 single guide RNA are linked to one another by intervening nucleotides and the linker can have a length of from 3 to 20 nucleotides (nt) (e.g., from 3 to 15, 3 to 12, 3 to 10, 3 to 8, 3 to 6, 3 to 5, 3 to 4, 4 to 20, 4 to 15, 4 to 12, 4 to 10, 4 to 8, 4 to 6, or 4 to 5 nt). In some embodiments, the linker of a CasP6 single guide RNA can have a length of from 3 to 100 nucleotides (nt) (e.g., from 3 to 80, 3 to 50, 3 to 30, 3 to 25, 3 to 20, 3 to 15, 3 to 12, 3 to 10, 3 to 8, 3 to 6, 3 to 5, 3 to 4, 4 to 100, 4 to 80, 4 to 50, 4 to 30, 4 to 25, 4 to 20, 4 to 15, 4 to 12, 4 to 10, 4 to 8, 4 to 6, or 4 to 5 nt). In some embodiments, the linker of a CasP6 single guide RNA can have a length of from 3 to 10 nucleotides (nt) (e.g., from 3 to 9, 3 to 8, 3 to 7, 3 to 6, 3 to 5, 3 to 4, 4 to 10, 4 to 9, 4 to 8, 4 to 7, 4 to 6, or 4 to 5 nt).

[0098] The targeting segment of a subject CasP6 guide RNA includes a guide sequence (i.e., a targeting sequence), which is a nucleotide sequence that is complementary to a sequence (a target site) in a target nucleic acid. In other words, the targeting segment of a CasP6 guide RNA can interact with a target nucleic acid (e.g., double stranded DNA (dsDNA), single stranded DNA (ssDNA), single stranded RNA (ssRNA), or double stranded RNA (dsRNA)) in a sequence-specific manner via hybridization (i.e., base pairing). The guide sequence of a CasP6 guide RNA can be modified (e.g., by genetic engineering)/designed to hybridize to any desired target sequence (e.g., while taking the PAM into account, e.g., when targeting a dsDNA target) within a target nucleic acid (e.g., a eukaryotic target nucleic acid such as genomic DNA or a target RNA such as an mRNA or viral RNA genome).

[0099] In some embodiments, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 60% or more (e.g., 65% or more, 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100%). In some cases, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 80% or more (e.g., 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100%). In some cases, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 90% or more (e.g., 95% or more, 97% or more, 98% or more, 99% or more, or 100%). In some cases, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 100%.

[0100] In some cases, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 100% over the seven contiguous 3 '-most nucleotides of the target site of the target nucleic acid. In some cases, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 60% or more (e.g., 70% or more, 75% or more, 80% or more, 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100%) over 19 or more (e.g., 20 or more, 21 or more, 22 or more, 23 or more, 24 or more, 25 or more) contiguous nucleotides. In some cases, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 80% or more (e.g., 85% or more, 90% or more, 95% or more, 97% or more, 98% or more, 99% or more, or 100%) over 19 or more (e.g., 20 or more, 21 or more, 22 or more, 23 or more, 24 or more, 25 or more) contiguous nucleotides. In some cases, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 90% or more (e.g., 95% or more, 97% or more, 98% or more, 99% or more, or 100%) over 19 or more (e.g., 20 or more, 21 or more, 22 or more, 23 or more, 24 or more, 25 or more) contiguous nucleotides. In some cases, the percent complementarity between the guide sequence and the target site of the target nucleic acid is 100% over 19 or more (e.g., 20 or more, 21 or more, 22 or more, 23 or more, 24 or more, 25 or more) contiguous nucleotides.

[0101] In some cases, the guide sequence has a length in a range of from 19-30 nucleotides (e.g., from 19-25, 19-22, 19-20, 20-30, 20-25, or 20-22 nt). In some cases, the guide sequence has a length in a range of from 19-25 nucleotides (e.g., from 19-22, 19-20, 20-25, 20-25, or 20-22 nt). In some cases, the guide sequence has a length of 19 or more nt (e.g., 20 or more, 21 or more, or 22 or more nt; 19 nt, 20 nt, 21 nt, 22 nt, 23 nt, 24 nt, 25 nt, etc.). In some cases the guide sequence has a length of 17 nt. In some cases the guide sequence has a length of 18 nt. In some cases the guide sequence has a length of 19 nt. In some cases the guide sequence has a length of

20 nt. In some cases the guide sequence has a length of 21 nt. In some cases the guide sequence has a length of 22 nt. In some cases the guide sequence has a length of 23 nt.

The protein-binding segment of a subject CasP6 guide RNA interacts with a CasP6 protein. The CasP6 guide RNA guides the bound CasP6 protein to a specific nucleotide sequence within target nucleic acid via the above mentioned guide sequence. In some embodiments, the protein-binding segment of a CasP6 guide RNA comprises two stretches of nucleotides (the duplex-forming segment of the activator and the duplex- forming segment of the targeter) that are complementary to one another and hybridize to form a double stranded RNA duplex (dsRNA duplex). Thus, the protein-binding segment includes a dsRNA duplex.

[0102] The duplex region of a subject CasP6 guide RNA (in dual guide or single guide RNA format) can include one or more (1, 2, 3, 4, 5, etc) mutations relative to a naturally occurring duplex region. For example, in some cases a base pair can be maintained while the nucleotides contributing to the base pair from each segment (targeter and activator) can be different. In some cases, the duplex region of a subject CasP6 guide RNA includes more paired bases, less paired bases, a smaller bulge, a larger bulge, fewer bulges, more bulges, or any convenient combination thereof, as compared to a naturally occurring duplex region (of a naturally occurring CasP6 guide RNA).

[0103] In some cases, the activator (e.g., activator-RNA) of a subject CasP6 guide RNA (in single guide RNA format) includes at least two internal RNA duplexes (i.e., two internal hairpins in addition to the activator/targeter dsRNA). The internal RNA duplexes (hairpins) of the activator can be positioned 5' of the activator/targeter dsRNA duplex. In some cases, the activator includes one hairpin positioned 5' of the activator/targeter dsRNA duplex. In some cases, the activator includes two hairpins positioned 5' of the activator/targeter dsRNA duplex. In some cases, the activator includes three hairpins positioned 5' of the activator/targeter dsRNA duplex. In some cases, the activator includes two or more hairpins (e.g., 3 or more or 4 or more hairpins) positioned 5' of the activator/targeter dsRNA duplex. In some cases, the activator includes 2 to 5 hairpins (e.g., 2 to 4, or 2 to 3 hairpins) positioned 5' of the activator/targeter dsRNA duplex.

[0104] In some cases, the activator-RNA (e.g., in single guide RNA format) comprises at least 2 nucleotides (nt) (e.g., at least 3 or at least 4 nt) 5' of the 5'-most hairpin stem. In some cases, the activator-RNA (e.g., in single guide RNA format) comprises at least 4 nt 5' of the 5'-most hairpin stem. In some cases, the activator-RNA (e.g., in single guide format) has a length of 65 nucleotides (nt) or more (e.g., 66 or more, 67 or more, 68 or more, 69 or more, 70 or more, or 75 or more nt). In some cases, the activator-RNA (e.g., in single guide format) has a length of 66 nt or more (e.g., 67 or more, 68 or more, 69 or more, 70 or more, or 75 or more nt). In some cases,

the activator- RNA (e.g., in single guide format) has a length of 67 nt or more (e.g., 68 or more, 69 or more, 70 or more, or 75 or more nt).

[0105] In some cases, the activator-RNA (e.g., in single or dual guide format) includes 45 or more nucleotides (nt) (e.g., 46 or more, 47 or more, 48 or more, 49 or more, 50 or more, 51 or more, 52 or more, 53 or more, 54 or more, or 55 or more nt) 5' of the dsRNA duplex formed between the activator and the targeter (the activator/targeter dsRNA duplex). In some cases, the activator is truncated at the 5' end relative to a naturally occurring CasP6 activator. In some cases, the activator is extended at the 5' end relative to a naturally occurring CasP6 activator.

[0106] Examples of various Cas9 guide RNAs can be found in the art, and in some cases variations similar to those introduced into Cas9 guide RNAs can also be introduced into CasP6 guide RNAs of the present disclosure. For example, see Jinek et al., Science. 2012 Aug 17;337(6096):816-21; Chylinski et al., RNA Biol. 2013 May;10(5):726-37; Ma et al., Biomed Res Int. 2013:270805; Hou et al., Proc Natl Acad Sci U S A. 2013 Sep 24;110(39): 15644-9; Jinek et al., Elife. 2013;2:e00471; Pattanayak et al., Nat Biotechnol. 2013 Sep;31(9):839-43; Qi et al, Cell. 2013 Feb 28;152(5): 1173-83; Wang et al., Cell. 2013 May 9;153(4):910-8; Auer et. al., Genome Res. 2013 Oct 31 ; Chen et. al., Nucleic Acids Res. 2013 Nov;41(20):e193; Cheng et. al., Cell Res. 2013 Oct;23(10): 1163-71 ; Cho et. al., Genetics. 2013 Nov;195(3): 1177-80; DiCarlo et al., Nucleic Acids Res. 2013 Apr;41(7):4336-43; Dickinson et. al., Nat Methods. 2013 Oct;10(10): 1028-34; Ebina et. al., Sci Rep. 2013;3:2510; Fujii et. al, Nucleic Acids Res. 2013 Nov 1;41(20):el87; Hu et. al., Cell Res. 2013 Nov 23(11): 1322-5; Jiang et. al., Nucleic Acids Res. 2013 Nov 1;41(20):e188; Larson et. al., Nat Protoc. 2013 Nov;8(11):2180-96; Mali et. at., Nat Methods. 2013 Oct;10(10):957-63; Nakayama et. al., Genesis. 2013 Dec;51(12):835-43; Ran et. al., Nat Protoc. 2013 Nov:8(11):2281-2308; Ran et. al., Cell. 2013 Sep 12;154(6): 1380-9; Upadhyay et. al., G3 (Bethesda). 2013 Dec 9;3(12):2233-8; Walsh et. al., Proc Natl Acad Sci U S A. 2013 Sep 24;110(39): 15514-5; Xie et. al., Mol Plant. 2013 Oct 9; Yang et. al., Cell. 2013 Sep 12;154(6): 1370-9; Briner et al., Mol Cell. 2014 Oct 23;56(2):333-9; and U.S. patents and patent applications: 8,906,616; 8,895,308; 8,889,418; 8,889,356; 8,871,445; 8,865,406; 8,795,965; 8,771,945; 8,697,359; 20140068797; 20140170753; 20140179006; 20140179770; 20140186843; 20140186919; 20140186958; 20140189896; 20140227787; 20140234972; 20140242664; 20140242699; 20140242700; 20140242702; 20140248702; 20140256046; 20140273037; 20140273226; 20140273230; 20140273231 ; 20140273232; 20140273233; 20140273234; 20140273235; 20140287938; 20140295556; 20140295557; 20140298547; 20140304853; 20140309487; 20140310828; 20140310830; 20140315985; 20140335063; 20140335620; 20140342456; 20140342457; 20140342458; 20140349400; 20140349405; 20140356867;

20140356956; 20140356958; 20140356959; 20140357523; 20140357530; 20140364333; and 20140377868; all of which are hereby incorporated by reference in their entirety.

[0107] In some cases, the term "activator" or "activator RNA" is used herein to mean a tracrRNA-like molecule (tracrRNA : "trans-acting CRISPR RNA") of a CasP6 dual guide RNA (and therefore of a CasP6 single guide RNA when the "activator" and the "targeter" are linked together by, e.g., intervening nucleotides). Thus, for example, a CasP6 guide RNA (dgRNA or sgRNA) comprises an activator sequence (e.g., a tracrRNA sequence). A tracr molecule (a tracrRNA) is a naturally existing molecule that hybridizes with a CRISPR RNA molecule (a crRNA) to form a CasP6 dual guide RNA. The term "activator" is used herein to encompass naturally existing tracrRNAs, but also to encompass tracrRNAs with modifications (e.g., truncations, extensions, sequence variations, base modifications, backbone modifications, linkage modifications, etc.) where the activator retains at least one function of a tracrRNA (e.g., contributes to the dsRNA duplex to which CasP6 protein binds). In some cases the activator provides one or more stem loops that can interact with CasP6 protein. An activator can be referred to as having a tracr sequence (tracrRNA sequence) and in some cases is a tracrRNA, but the term "activator" is not limited to naturally existing tracrRNAs.

[0108] In some cases (e.g., in some cases where the guide RNA is in single guide format), the activator-RNA is truncated (shorter) relative to the corresponding wild type tracrRNA. In some cases (e.g., in some cases where the guide RNA is in single guide format) the activator-RNA is not truncated (shorter) relative to the corresponding wild type tracrRNA. In some cases (e.g., in some cases where the guide RNA is in single guide format) the activator-RNA has a length that is greater than 50 nt (e.g., greater than 55nt, greater than 60nt, greater than 65 nt, greater than 70nt, greater than 75nt, greater than 80nt). In some cases (e.g., in some cases where the guide RNA is in single guide format) the activator- RNA has a length that is greater than 80 nt. In some cases (e.g., in some cases where the guide RNA is in single guide format) the activator-RNA has a length in a range of from 51 to 90 nt (e.g., from 51-85, 51- 84, 55-90, 55-85, 55-84, 60-90, 60-85, 60-84, 65-90, 65-85, 65-84, 70-90, 70-85, 70-84, 75-90, 75-85, 75- 84, 80-90, 80-85, or 80-84 nt). In some cases (e.g., in some cases where the guide RNA is in single guide format) the activator-RNA has a length in a range of from 80-90 nt. In certain embodiments, the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6 or comprises an RNA molecule of SEQ ID NO:7, SEQ ID NO: 8, or SEQ ID NO: 9. In certain embodiments, the guide RNAs can comprise variants of SEQ ID NO:7, SEQ ID NO: 8, or SEQ ID NO: 9 having at least 90%, 95%, 97%, 98%, 99%, or 99.5% sequence identity to SEQ ID NO:7, SEQ ID NO: 8, or SEQ ID NO: 9. In certain embodiments,

the guide RNAs can comprise variants of SEQ ID NO:7, SEQ ID NO: 8, or SEQ ID NO: 9 having at least one, two, three, or four nucleotide substitutions, deletions, and/or insertions. Also provided are DNA molecules encoding any of the aforementioned RNA molecules, including DNA molecules comprising SEQ ID NO:4, SEQ ID NO: 5, or SEQ ID NO: 6 as well as variants thereof. In certain embodiments, the DNA variants encoding the variant guide RNAs can have at least 90%, 95%, 97%, 98%, 99%, or 99.5% sequence identity to SEQ ID NO:4, SEQ ID NO: 5, or SEQ ID NO: 6. In certain embodiments, the DNA variants encoding the variant guide RNAs can have at least one, two, three, or four nucleotide substitutions, deletions, and/or insertions in SEQ ID NO: 4, SEQ ID NO: 5, or SEQ ID NO: 6.

[0109] The term "targeter" or "targeter RNA" is used herein to refer to a crRNA-like molecule (crRNA: "CRISPR RNA") of a CasP6 dual guide RNA (and therefore of a CasP6 single guide RNA when the "activator" and the "targeter" are linked together, e.g., by intervening nucleotides). Thus, for example, a CasP6 guide RNA (dgRNA or sgRNA) comprises a guide sequences and a duplex-forming segment (e.g., a duplex forming segment of a crRNA, which can also be referred to as a crRNA repeat). Because the sequence of a targeting segment (the segment that hybridizes with a target sequence of a target nucleic acid) of a targeter is modified by a user to hybridize with a desired target nucleic acid, the sequence of a targeter will often be a non-naturally occurring sequence. However, the duplex-forming segment of a targeter (described in more detail herein), which hybridizes with the duplex-forming segment of an activator, can include a naturally existing sequence (e.g., can include the sequence of a duplex-forming segment of a naturally existing crRNA, which can also be referred to as a crRNA repeat). Thus, the term targeter is used herein to distinguish from naturally occurring crRNAs, despite the fact that part of a targeter (e.g., the duplex-forming segment) often includes a naturally occurring sequence from a crRNA. However, the term "targeter" encompasses naturally occurring crRNAs.

[0110] As noted above, a targeter comprises both the guide sequence of the CasP6 guide RNA and a stretch (a "duplex-forming segment") of nucleotides that forms one half of the dsRNA duplex of the protein-binding segment of the CasP6 guide RNA. A corresponding tracrRNA-like molecule (activator) comprises a stretch of nucleotides (a duplex-forming segment) that forms the other half of the dsRNA duplex of the protein-binding segment of the CasP6 guide RNA. In other words, a stretch of nucleotides of the targeter is complementary to and hybridizes with a stretch of nucleotides of the activator to form the dsRNA duplex of the protein-binding segment of a CasP6 guide RNA. As such, each targeter can be said to have a corresponding activator (which has a region that hybridizes with the targeter). The targeter molecule additionally provides the guide sequence. Thus, a targeter and an activator (as a corresponding pair) hybridize to form a

CasP6 guide RNA. The particular sequence of a given naturally existing crRNA or tracrRNA molecule can be characteristic of the species in which the RNA molecules are found.

[0111] The present disclosure provides an engineered, non-naturally occurring CasP6 system. A CasP6 system of the present disclosure can comprise: a) a CasP6 polypeptide of the present disclosure and a CasP6 guide RNA; b) a CasP6 polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; c) a CasP6 fusion polypeptide of the present disclosure and a CasP6 guide RNA; d) a CasP6 fusion polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; e) an mRNA encoding a CasP6 polypeptide of the present disclosure; and a CasP6 guide RNA; f) an mRNA encoding a CasP6 polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; g) an mRNA encoding a CasP6 fusion polypeptide of the present disclosure; and a CasP6 guide RNA; h) an mRNA encoding a CasP6 fusion polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; i) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure and a nucleotide sequence encoding a CasP6 guide RNA; j) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, a nucleotide sequence encoding a CasP6 guide RNA, and a nucleotide sequence encoding a donor template nucleic acid; k) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure and a nucleotide sequence encoding a CasP6 guide RNA; 1) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, a nucleotide sequence encoding a CasP6 guide RNA, and a nucleotide sequence encoding a donor template nucleic acid; m) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; n) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; and a donor template nucleic acid; o) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; p) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; and a donor template nucleic acid; q) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6

polypeptide of the present disclosure, a nucleotide sequence encoding a first CasP6 guide RNA, and a nucleotide sequence encoding a second CasP6 guide RNA; or r) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, a nucleotide sequence encoding a first CasP6 guide RNA, and a nucleotide sequence encoding a second CasP6 guide RNA; or some variation of one of (a) through (r).

[0112] The present disclosure provides one or more nucleic acids comprising one or more of: a donor polynucleotide sequence, a nucleotide sequence encoding a CasP6 polypeptide (e.g., a wild type CasP6 protein, a nickase CasP6 protein, a dCasP6 protein, chimeric CasP6 protein, and the like), a CasP6 guide RNA, and a nucleotide sequence encoding a CasP6 guide RNA (which can include two separate nucleotide sequences in the case of dual guide RNA format or which can include a single nucleotide sequence in the case of single guide RNA format). The present disclosure provides a nucleic acid comprising a nucleotide sequence encoding a CasP6 fusion polypeptide. The present disclosure provides a recombinant expression vector that comprises a nucleotide sequence encoding a CasP6 polypeptide. The present disclosure provides a recombinant expression vector that comprises a nucleotide sequence encoding a CasP6 fusion polypeptide. The present disclosure provides a recombinant expression vector that comprises: a) a nucleotide sequence encoding a CasP6 polypeptide; and b) a nucleotide sequence encoding a CasP6 guide RNA(s). The present disclosure provides a recombinant expression vector that comprises: a) a nucleotide sequence encoding a CasP6 fusion polypeptide; and b) a nucleotide sequence encoding a CasP6 guide RNA(s). In some cases, the nucleotide sequence encoding the CasP6 protein and/or the nucleotide sequence encoding the CasP6 guide RNA is operably linked to a promoter that is operable in a cell type of choice (e.g., a prokarytoic cell, a eukaryotic cell, a plant cell, an animal cell, a mammalian cell, a primate cell, a rodent cell, a human cell, etc.).

[0113] In some cases, a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure is codon optimized. This type of optimization can entail a mutation of a CasP6 -encoding nucleotide sequence to mimic the codon preferences of the intended host organism or cell while encoding the same protein. Thus, the codons can be changed, but the encoded protein remains unchanged. For example, if the intended target cell was a human cell, a human codon-optimized CasP6-encoding nucleotide sequence could be used. As another non-limiting example, if the intended host cell were a mouse cell, then a mouse codon-optimized CasP6-encoding nucleotide sequence could be generated. As another non-limiting example, if the intended host cell were a plant cell, then a plant codon-optimized CasP6-encoding nucleotide sequence could be generated. As another non-limiting example, if the intended host cell were an insect cell, then an insect codon-optimized CasP6-encoding nucleotide sequence could be generated.

[0114] The present disclosure provides one or more recombinant expression vectors that include (in different recombinant expression vectors in some cases, and in the same recombinant expression vector in some cases): (i) a nucleotide sequence of a donor template nucleic acid (where the donor template comprises a nucleotide sequence having homology to a target sequence of a target nucleic acid (e.g., a target genome)); (ii) a nucleotide sequence that encodes a CasP6 guide RNA that hybridizes to a target sequence of the target locus of the targeted genome (e.g., a single or dual guide RNA) (e.g., operably linked to a promoter that is operable in a target cell such as a eukaryotic cell); and (iii) a nucleotide sequence encoding a CasP6 protein (e.g., operably linked to a promoter that is operable in a target cell such as a eukaryotic cell). The present disclosure provides one or more recombinant expression vectors that include (in different recombinant expression vectors in some cases, and in the same recombinant expression vector in some cases): (i) a nucleotide sequence of a donor template nucleic acid (where the donor template comprises a nucleotide sequence having homology to a target sequence of a target nucleic acid (e.g., a target genome)); and (ii) a nucleotide sequence that encodes a CasP6 guide RNA that hybridizes to a target sequence of the target locus of the targeted genome (e.g., a single or dual guide RNA) (e.g., operably linked to a promoter that is operable in a target cell such as a eukaryotic cell). The present disclosure provides one or more recombinant expression vectors that include (in different recombinant expression vectors in some cases, and in the same recombinant expression vector in some cases): (i) a nucleotide sequence that encodes a CasP6 guide RNA that hybridizes to a target sequence of the target locus of the targeted genome or target RNA (e.g., a single or dual guide RNA) (e.g., operably linked to a promoter that is operable in a target cell such as a eukaryotic cell); and (ii) a nucleotide sequence encoding a CasP6 protein (e.g., operably linked to a promoter that is operable in a target cell such as a eukaryotic cell).

[0115] Suitable expression vectors include viral expression vectors (e.g. viral vectors based on vaccinia virus; poliovirus; adenovirus (see, e.g., Li et al., Invest Opthalmol Vis Sci 35:2543 2549, 1994; Borras et al, Gene Ther 6:515 524, 1999; Li and Davidson, PNAS 92:7700 7704, 1995; Sakamoto et al., H Gene Ther 5: 1088 1097, 1999; WO 94/12649, WO 93/03769; WO 93/19191 ; WO 94/28938; WO 95/11984 and WO 95/00655); adeno-associated virus (AAV) (see, e.g., Ali et al., Hum Gene Ther 9:81 86, 1998, Flannery et al., PNAS 94:6916 6921, 1997; Bennett et al., Invest Opthalmol Vis Sci 38:2857 2863, 1997; Jomary et al., Gene Ther 4:683 690, 1997, Rolling et al., Hum Gene Ther 10:641 648, 1999; Ali et al., Hum Mol Genet 5:591 594, 1996; Srivastava in WO 93/09239, Samulski et al., J. Vir. (1989) 63:3822-3828; Mendelson et al., Virol. (1988) 166:154-165; and Flotte et al., PNAS (1993) 90: 10613-10617); SV40; herpes simplex virus; human immunodeficiency virus (see, e.g., Miyoshi et al., PNAS 94: 10319 23, 1997; Takahashi

et al., J Virol 73:7812 7816, 1999); a retroviral vector (e.g., Murine Leukemia Virus, spleen necrosis virus, and vectors derived from retroviruses such as Rous Sarcoma Virus, Harvey Sarcoma Virus, avian leukosis virus, a lentivirus, human immunodeficiency virus, myeloproliferative sarcoma virus, and mammary tumor virus); and the like. In some cases, a recombinant expression vector of the present disclosure is a recombinant adeno-associated virus (AAV) vector. In some cases, a recombinant expression vector of the present disclosure is a recombinant lentivirus vector. In some cases, a recombinant expression vector of the present disclosure is a recombinant retroviral vector.

[0116] Depending on the host/vector system utilized, any of a number of suitable transcription and translation control elements, including constitutive and inducible promoters, transcription enhancer elements, transcription terminators, etc. may be used in the expression vector.

[0117] In some embodiments, a nucleotide sequence encoding a CasP6 guide RNA is operably linked to a control element, e.g., a transcriptional control element, such as a promoter. In some embodiments, a nucleotide sequence encoding a CasP6 protein or a CasP6 fusion polypeptide is operably linked to a control element, e.g., a transcriptional control element, such as a promoter.

[0118] The transcriptional control element can be a promoter. In some cases, the promoter is a constitutively active promoter. In some cases, the promoter is a regulatable promoter. In some cases, the promoter is an inducible promoter. In some cases, the promoter is a tissue-specific promoter. In some cases, the promoter is a cell type-specific promoter. In some cases, the transcriptional control element (e.g., the promoter) is functional in a targeted cell type or targeted cell population. For example, in some cases, the transcriptional control element can be functional in eukaryotic cells, e.g., hematopoietic stem cells (e.g., mobilized peripheral blood (mPB) CD34(+) cell, bone marrow (BM) CD34(+) cell, etc.).

[0119] Nonlimiting examples of eukaryotic promoters (promoters functional in a eukaryotic cell) include EF1a, those from cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, early and late SV40, long terminal repeats (LTRs) from retrovirus, and mouse metallothionein-I. Selection of the appropriate vector and promoter is well within the level of ordinary skill in the art. The expression vector may also contain a ribosome binding site for translation initiation and a transcription terminator. The expression vector may also include appropriate sequences for amplifying expression. The expression vector may also include nucleotide sequences encoding protein tags (e.g., 6xHis tag, hemagglutinin tag, fluorescent protein, etc.) that can be fused to the CasP6 protein, thus resulting in a chimeric CasP6 polypeptide.

[0120] In some embodiments, a nucleotide sequence encoding a CasP6 guide RNA and/or a CasP6 fusion polypeptide is operably linked to an inducible promoter. In some embodiments, a nucleotide sequence encoding a CasP6 guide RNA and/or a CasP6 fusion protein is operably linked to a constitutive promoter.

[0121] A promoter can be a constitutively active promoter (i.e., a promoter that is constitutively in an active/"ON" state), it may be an inducible promoter (i.e., a promoter whose state, active/"ON" or inactive/"OFF", is controlled by an external stimulus, e.g., the presence of a particular temperature, compound, or protein.), it may be a spatially restricted promoter (i.e., transcriptional control element, enhancer, etc.)(e.g., tissue specific promoter, cell type specific promoter, etc.), and it may be a temporally restricted promoter (i.e., the promoter is in the "ON" state or "OFF" state during specific stages of embryonic development or during specific stages of a biological process, e.g., hair follicle cycle in mice).

[0122] Suitable promoters can be derived from viruses and can therefore be referred to as viral promoters, or they can be derived from any organism, including prokaryotic or eukaryotic organisms. Suitable promoters can be used to drive expression by any RNA polymerase (e.g., pol I, pol II, pol III). Exemplary promoters include, but are not limited to the SV40 early promoter, mouse mammary tumor virus long terminal repeat (LTR) promoter; adenovirus major late promoter (Ad MLP); a herpes simplex virus (HSV) promoter, a cytomegalovirus (CMV) promoter such as the CMV immediate early promoter region (CMVIE), a rous sarcoma virus (RSV) promoter, a human U6 small nuclear promoter (U6) (Miyagishi et al., Nature Biotechnology 20, 497 - 500 (2002)), an enhanced U6 promoter (e.g., Xia et al., Nucleic Acids Res. 2003 Sep 1 ;31(17)), a human HI promoter (HI), and the like.

[0123] In some cases, a nucleotide sequence encoding a CasP6 guide RNA is operably linked to (under the control of) a promoter operable in a eukaryotic cell (e.g., a U6 promoter, an enhanced U6 promoter, an HI promoter, and the like). As would be understood by one of ordinary skill in the art, when expressing an RNA (e.g., a guide RNA) from a nucleic acid (e.g., an expression vector) using a U6 promoter (e.g., in a eukaryotic cell), or another PolIII promoter, the RNA may need to be mutated if there are several Ts in a row (coding for Us in the RNA). This is because a string of Ts (e.g., 5 Ts) in DNA can act as a terminator for polymerase III (PolIII). Thus, in order to ensure transcription of a guide RNA (e.g., the activator portion and/or targeter portion, in dual guide or single guide format) in a eukaryotic cell it may sometimes be necessary to modify the sequence encoding the guide RNA to eliminate runs of Ts. In some cases, a nucleotide sequence encoding a CasP6 protein (e.g., a wild type CasP6 protein, a nickase CasP6 protein, a dCasP6 protein, a chimeric CasP6 protein and the like) is operably linked to a promoter operable in a

eukaryotic cell (e.g., a CMV promoter, an EFla promoter, an estrogen receptor-regulated promoter, and the like).

[0124] Examples of inducible promoters include, but are not limited toT7 RNA polymerase promoter, T3 RNA polymerase promoter, Isopropyl-beta-D-thiogalactopyranoside (IPTG) -regulated promoter, lactose induced promoter, heat shock promoter, Tetracycline-regulated promoter, Steroid- regulated promoter, Metal-regulated promoter, estrogen receptor-regulated promoter, etc. Inducible promoters can therefore be regulated by molecules including, but not limited to, doxycycline; estrogen and/or an estrogen analog; IPTG; etc.

[0125] Inducible promoters suitable for use include any inducible promoter described herein or known to one of ordinary skill in the art. Examples of inducible promoters include, without limitation, chemically/biochemically-regulated and physically-regulated promoters such as alcohol-regulated promoters, tetracycline-regulated promoters (e.g., anhydrotetracycline (aTc)-responsive promoters and other tetracycline -responsive promoter systems, which include a tetracycline repressor protein (tetR), a tetracycline operator sequence (tetO) and a tetracycline transactivator fusion protein (tTA)), steroid- regulated promoters (e.g., promoters based on the rat glucocorticoid receptor, human estrogen receptor, moth ecdysone receptors, and promoters from the steroid/retinoid/thyroid receptor superfamily), metal- regulated promoters (e.g., promoters derived from metallothionein (proteins that bind and sequester metal ions) genes from yeast, mouse and human), pathogenesis-regulated promoters (e.g., induced by salicylic acid, ethylene or benzothiadiazole (BTH)), temperature/heat-inducible promoters (e.g., heat shock promoters), and light-regulated promoters (e.g., light responsive promoters from plant cells).

[0126] In some cases, the promoter is a spatially restricted promoter (i.e., cell type specific promoter, tissue specific promoter, etc.) such that in a multi-cellular organism, the promoter is active (i.e., "ON") in a subset of specific cells. Spatially restricted promoters may also be referred to as enhancers, transcriptional control elements, control sequences, etc. Any convenient spatially restricted promoter may be used as long as the promoter is functional in the targeted host cell (e.g., eukaryotic cell; prokaryotic cell).

[0127] In some cases, the promoter is a reversible promoter. Suitable reversible promoters, including reversible inducible promoters are known in the art. Such reversible promoters may be isolated and derived from many organisms, e.g., eukaryotes and prokaryotes. Modification of reversible promoters derived from a first organism for use in a second organism, e.g., a first prokaryote and a second a eukaryote, a first eukaryote and a second a prokaryote, etc., is well known in the art. Such reversible promoters, and systems based on such reversible promoters but also comprising additional control proteins, include, but are not limited to, alcohol regulated

promoters (e.g., alcohol dehydrogenase I (alcA) gene promoter, promoters responsive to alcohol transactivator proteins (AlcR), etc.), tetracycline regulated promoters, (e.g., promoter systems including Tet Activators, TetON, TetOFF, etc.), steroid regulated promoters (e.g., rat glucocorticoid receptor promoter systems, human estrogen receptor promoter systems, retinoid promoter systems, thyroid promoter systems, ecdysone promoter systems, mifepristone promoter systems, etc.), metal regulated promoters (e.g., metallothionein promoter systems, etc.), pathogenesis-related regulated promoters (e.g., salicylic acid regulated promoters, ethylene regulated promoters, benzothiadiazole regulated promoters, etc.), temperature regulated promoters (e.g., heat shock inducible promoters (e.g., HSP-70, HSP-90, soybean heat shock promoter, etc.), light regulated promoters, synthetic inducible promoters, and the like.

[0128] Methods of introducing a nucleic acid (e.g., a nucleic acid comprising a donor polynucleotide sequence, one or more nucleic acids encoding a CasP6 protein and/or a CasP6 guide RNA, and the like) into a host cell are known in the art, and any convenient method can be used to introduce a nucleic acid (e.g., an expression construct) into a cell. Suitable methods include e.g., viral infection, transfection, lipofection, electroporation, calcium phosphate precipitation, polyethyleneimine (PEI)- mediated transfection, DEAE-dextran mediated transfection, liposome-mediated transfection, particle gun technology, calcium phosphate precipitation, direct microinjection, nanoparticle -mediated nucleic acid delivery, and the like.

[0129] Introducing the recombinant expression vector into cells can occur in any culture media and under any culture conditions that promote the survival of the cells. Introducing the recombinant expression vector into a target cell can be carried out in vivo or ex vivo. Introducing the recombinant expression vector into a target cell can be carried out in vitro.

[0130] In some embodiments, a CasP6 protein can be provided as RNA. The RNA can be provided by direct chemical synthesis or may be transcribed in vitro from a DNA (e.g., encoding the CasP6 protein). Once synthesized, the RNA may be introduced into a cell by any of the well-known techniques for introducing nucleic acids into cells (e.g., microinjection, electroporation, transfection, etc.).

[0131] Nucleic acids may be provided to the cells using well-developed transfection techniques; see, e.g. Angel and Yanik (2010) PLoS ONE 5(7): el 1756, and the commercially available TransMessenger® reagents from Qiagen, Stemfect™ RNA Transfection Kit from Stemgent, and TransIT®-mRNA Transfection Kit from Mirus Bio LLC. See also Beumer et al. (2008) PNAS 105(50): 19821-19826.

[0132] Vectors may be provided directly to a target host cell. In other words, the cells are contacted with vectors comprising the subject nucleic acids (e.g., recombinant expression vectors

having the donor template sequence and encoding the CasP6 guide RNA; recombinant expression vectors encoding the CasP6 protein; etc.) such that the vectors are taken up by the cells. Methods for contacting cells with nucleic acid vectors that are plasmids, include electroporation, calcium chloride transfection, microinjection, and lipofection are well known in the art. For viral vector delivery, cells can be contacted with viral particles comprising the subject viral expression vectors.

[0133] Retroviruses, for example, lentiviruses, are suitable for use in methods of the present disclosure. Commonly used retroviral vectors are "defective", i.e. unable to produce viral proteins required for productive infection. Rather, replication of the vector requires growth in a packaging cell line. To generate viral particles comprising nucleic acids of interest, the retroviral nucleic acids comprising the nucleic acid are packaged into viral capsids by a packaging cell line. Different packaging cell lines provide a different envelope protein (ecotropic, amphotropic or xenotropic) to be incorporated into the capsid, this envelope protein determining the specificity of the viral particle for the cells (ecotropic for murine and rat; amphotropic for most mammalian cell types including human, dog and mouse; and xenotropic for most mammalian cell types except murine cells). The appropriate packaging cell line may be used to ensure that the cells are targeted by the packaged viral particles. Methods of introducing subject vector expression vectors into packaging cell lines and of collecting the viral particles that are generated by the packaging lines are well known in the art. Nucleic acids can also introduced by direct micro-injection (e.g., injection of RNA).

[0134] Vectors used for providing the nucleic acids encoding CasP6 guide RNA and/or a CasP6 polypeptide to a target host cell can include suitable promoters for driving the expression, that is, transcriptional activation, of the nucleic acid of interest. In other words, in some cases, the nucleic acid of interest will be operably linked to a promoter. This may include ubiquitously acting promoters, for example, the CMV- -actin promoter, or inducible promoters, such as promoters that are active in particular cell populations or that respond to the presence of drugs such as tetracycline. By transcriptional activation, it is intended that transcription will be increased above basal levels in the target cell by 10 fold, by 100 fold, more usually by 1000 fold. In addition, vectors used for providing a nucleic acid encoding a CasP6 guide RNA and/or a CasP6 protein to a cell may include nucleic acid sequences that encode for selectable markers in the target cells, so as to identify cells that have taken up the CasP6 guide RNA and/or CasP6 protein.

[0135] A nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide, or a CasP6 fusion polypeptide, is in some cases an RNA. Thus, a CasP6 fusion protein can be

introduced into cells as RNA. Methods of introducing RNA into cells are known in the art and may include, for example, direct injection, transfection, or any other method used for the introduction of DNA. A CasP6 protein may instead be provided to cells as a polypeptide. Such a polypeptide may optionally be fused to a polypeptide domain that increases solubility of the product. The domain may be linked to the polypeptide through a defined protease cleavage site, e.g. a TEV sequence, which is cleaved by TEV protease. The linker may also include one or more flexible sequences, e.g. from 1 to 10 glycine residues. In some embodiments, the cleavage of the fusion protein is performed in a buffer that maintains solubility of the product, e.g. in the presence of from 0.5 to 2 M urea, in the presence of polypeptides and/or polynucleotides that increase solubility, and the like. Domains of interest include endosomolytic domains, e.g. influenza HA domain; and other polypeptides that aid in production, e.g. IF2 domain, GST domain, GRPE domain, and the like. The polypeptide may be formulated for improved stability. For example, the peptides may be PEGylated, where the polyethyleneoxy group provides for enhanced lifetime in the blood stream.

[0136] Additionally or alternatively, a CasP6 polypeptide of the present disclosure may be fused to a polypeptide permeant domain to promote uptake by the cell. A number of permeant domains are known in the art and may be used in the non-integrating polypeptides of the present disclosure, including peptides, peptidomimetics, and non-peptide carriers. For example, a permeant peptide may be derived from the third alpha helix of Drosophila melanogaster transcription factor Antennapaedia, referred to as penetratin, which comprises the amino acid sequence of SEQ ID NO: 78. As another example, the permeant peptide comprises the HIV-1 tat basic region amino acid sequence, which may include, for example, amino acids 49-57 of naturally-occurring tat protein. Other permeant domains include poly-arginine motifs, for example, the region of amino acids 34-56 of HIV-1 rev protein, nona- arginine, octa-arginine, and the like. (See, for example, Futaki et al. (2003) Curr Protein Pept Sci. 2003 Apr; 4(2): 87-9 and 446; and Wender et al. (2000) Proc. Natl. Acad. Sci. U.S.A 2000 Nov. 21;97(24): 13003-8; published U.S. Patent applications 20030220334; 20030083256; 20030032593; and 20030022831, herein specifically incorporated by reference for the teachings of translocation peptides and peptoids). The nona-arginine (R9) sequence is one of the more efficient PTDs that have been characterized (Wender et al. 2000; Uemura et al. 2002). The site at which the fusion is made may be selected in order to optimize the biological activity, secretion or binding characteristics of the polypeptide. The optimal site will be determined by routine experimentation.

[0137] A CasP6 polypeptide of the present disclosure may be produced in vitro or by eukaryotic cells or by prokaryotic cells, and it may be further processed by unfolding, e.g. heat denaturation, dithiothreitol reduction, etc. and may be further refolded, using methods known in the art.

[0138] Modifications of interest that do not alter primary sequence include chemical derivatization of polypeptides, e.g., acylation, acetylation, carboxylation, amidation, etc. Also included are modifications of glycosylation, e.g. those made by modifying the glycosylation patterns of a polypeptide during its synthesis and processing or in further processing steps; e.g. by exposing the polypeptide to enzymes which affect glycosylation, such as mammalian glycosylating or deglycosylating enzymes. Also embraced are sequences that have phosphorylated amino acid residues, e.g. phosphotyrosine, phosphoserine, or phosphothreonine.

[0139] Also suitable for inclusion in embodiments of the present disclosure are nucleic acids (e.g., encoding a CasP6 guide RNA, encoding a CasP6 fusion protein, etc.) and proteins (e.g., a CasP6 fusion protein derived from a wild type protein or a variant protein) that have been modified using ordinary molecular biological techniques and synthetic chemistry so as to improve their resistance to proteolytic degradation, to change the target sequence specificity, to optimize solubility properties, to alter protein activity (e.g., transcription modulatory activity, enzymatic activity, etc.) or to render them more suitable. Analogs of such polypeptides include those containing residues other than naturally occurring L-amino acids, e.g. D-amino acids or non-naturally occurring synthetic amino acids. D-amino acids may be substituted for some or all of the amino acid residues.

[0140] A CasP6 polypeptide of the present disclosure may be prepared by in vitro synthesis, using conventional methods as known in the art. Various commercial synthetic apparatuses are available, for example, automated synthesizers by Applied Biosystems, Inc., Beckman, etc. By using synthesizers, naturally occurring amino acids may be substituted with unnatural amino acids. The particular sequence and the manner of preparation will be determined by convenience, economics, purity required, and the like.

[0141] If desired, various groups may be introduced into the peptide during synthesis or during expression, which allow for linking to other molecules or to a surface. Thus cysteines can be used to make thioethers, histidines for linking to a metal ion complex, carboxyl groups for forming amides or esters, amino groups for forming amides, and the like.

[0142] A CasP6 polypeptide of the present disclosure may also be isolated and purified in accordance with conventional methods of recombinant synthesis. A lysate may be prepared of the expression host and the lysate purified using high performance liquid chromatography (HPLC), exclusion chromatography, gel electrophoresis, affinity chromatography, or other purification

technique. For the most part, the compositions which are used will comprise 20% or more by weight of the desired product, more usually 75% or more by weight, preferably 95% or more by weight, and for therapeutic purposes, usually 99.5% or more by weight, in relation to contaminants related to the method of preparation of the product and its purification. Usually, the percentages will be based upon total protein. Thus, in some cases, a CasP6 polypeptide, or a CasP6 fusion polypeptide, of the present disclosure is at least 80% pure, at least 85% pure, at least 90% pure, at least 95% pure, at least 98% pure, or at least 99% pure (e.g., free of contaminants, non-CasP6 proteins or other macromolecules, etc.).

[0143] To induce cleavage or any desired modification to a target nucleic acid (e.g., genomic DNA or target RNA), or any desired modification to a polypeptide associated with target nucleic acid, the CasP6 guide RNA and/or the CasP6 polypeptide of the present disclosure and/or the donor template sequence, whether they be introduced as nucleic acids or polypeptides, are provided to the cells for about 30 minutes to about 24 hours, e.g., 1 hour, 1.5 hours, 2 hours, 2.5 hours, 3 hours, 3.5 hours 4 hours, 5 hours, 6 hours, 7 hours, 8 hours, 12 hours, 16 hours, 18 hours, 20 hours, or any other period from about 30 minutes to about 24 hours, which may be repeated with a frequency of about every day to about every 4 days, e.g., every 1.5 days, every 2 days, every 3 days, or any other frequency from about every day to about every four days. The agent(s) may be provided to the subject cells one or more times, e.g. one time, twice, three times, or more than three times, and the cells allowed to incubate with the agent(s) for some amount of time following each contacting event e.g. 16-24 hours, after which time the media is replaced with fresh media and the cells are cultured further.

[0144] In cases in which two or more different targeting complexes are provided to the cell (e.g., two different CasP6 guide RNAs that are complementary to different sequences within the same or different target nucleic acid), the complexes may be provided simultaneously (e.g. as two polypeptides and/or nucleic acids), or delivered simultaneously. Alternatively, they may be provided consecutively, e.g. the targeting complex being provided first, followed by the second targeting complex, etc. or vice versa.

[0145] To improve the delivery of a DNA vector into a target cell, the DNA can be protected from damage and its entry into the cell facilitated, for example, by using lipoplexes and polyplexes. Thus, in some cases, a nucleic acid of the present disclosure (e.g., a recombinant expression vector of the present disclosure) can be covered with lipids in an organized structure like a micelle or a liposome. When the organized structure is complexed with DNA it is called a lipoplex. There are three types of lipids, anionic (negatively-charged), neutral, or cationic (positively-charged). Lipoplexes that utilize cationic lipids have proven utility for gene transfer.

Cationic lipids, due to their positive charge, naturally complex with the negatively charged DNA. Also as a result of their charge, they interact with the cell membrane. Endocytosis of the lipoplex then occurs, and the DNA is released into the cytoplasm. The cationic lipids also protect against degradation of the DNA by the cell.

[0146] Complexes of polymers with DNA are called polyplexes. Most polyplexes consist of cationic polymers and their production is regulated by ionic interactions. One large difference between the methods of action of polyplexes and lipoplexes is that polyplexes cannot release their DNA load into the cytoplasm, so to this end, co-transfection with endosome -lytic agents (to lyse the endosome that is made during endocytosis) such as inactivated adenovirus must occur. However, this is not always the case; polymers such as polyethylenimine have their own method of endosome disruption as does chitosan and trimethylchitosan.

[0147] Dendrimers, a highly branched macromolecule with a spherical shape, may be also be used to genetically modify stem cells. The surface of the dendrimer particle may be functionalized to alter its properties. In particular, it is possible to construct a cationic dendrimer (i.e., one with a positive surface charge). When in the presence of genetic material such as a DNA plasmid, charge complementarity leads to a temporary association of the nucleic acid with the cationic dendrimer. On reaching its destination, the dendrimer-nucleic acid complex can be taken up into a cell by endocytosis.

[0148] In some cases, a nucleic acid of the disclosure (e.g., an expression vector) includes an insertion site for a guide sequence of interest. For example, a nucleic acid can include an insertion site for a guide sequence of interest, where the insertion site is immediately adjacent to a nucleotide sequence encoding the portion of a CasP6 guide RNA that does not change when the guide sequence is changed to hybridize to a desired target sequence (e.g., sequences that contribute to the CasP6 binding aspect of the guide RNA, e.g, the sequences that contribute to the dsRNA duplex(es) of the CasP6 guide RNA - this portion of the guide RNA can also be referred to as the 'scaffold' or 'constant region' of the guide RNA). Thus, in some cases, a subject nucleic acid (e.g., an expression vector) includes a nucleotide sequence encoding a CasP6 guide RNA, except that the portion encoding the guide sequence portion of the guide RNA is an insertion sequence (an insertion site). An insertion site is any nucleotide sequence used for the insertion of a the desired sequence. "Insertion sites" for use with various technologies are known to those of ordinary skill in the art and any convenient insertion site can be used. An insertion site can be for any method for manipulating nucleic acid sequences. For example, in some cases the insertion site is a multiple cloning site (MCS) (e.g., a site including one or more restriction enzyme recognition sequences), a site for ligation independent cloning, a site for recombination based

cloning (e.g., recombination based on att sites), a nucleotide sequence recognized by a CRISPR/Cas (e.g. Cas9) based technology, and the like.

[0149] An insertion site can be any desirable length, and can depend on the type of insertion site (e.g., can depend on whether (and how many) the site includes one or more restriction enzyme recognition sequences, whether the site includes a target site for a CRISPR/Cas protein, etc.). In some cases, an insertion site of a subject nucleic acid is 3 or more nucleotides (nt) in length (e.g., 5 or more, 8 or more, 10 or more, 15 or more, 17 or more, 18 or more, 19 or more, 20 or more or 25 or more, or 30 or more nt in length). In some cases, the length of an insertion site of a subject nucleic acid has a length in a range of from 2 to 50 nucleotides (nt) (e.g., from 2 to 40 nt, from 2 to 30 nt, from 2 to 25 nt, from 2 to 20 nt, from 5 to 50 nt, from 5 to 40 nt, from 5 to 30 nt, from 5 to 25 nt, from 5 to 20 nt, from 10 to 50 nt, from 10 to 40 nt, from 10 to 30 nt, from 10 to 25 nt, from 10 to 20 nt, from 17 to 50 nt, from 17 to 40 nt, from 17 to 30 nt, from 17 to 25 nt). In some cases, the length of an insertion site of a subject nucleic acid has a length in a range of from 5 to 40 nt.

[0150] In some embodiments, a subject nucleic acid (e.g., a CasP6 guide RNA) has one or more modifications, e.g., a base modification, a backbone modification, etc., to provide the nucleic acid with a new or enhanced feature (e.g., improved stability). A nucleoside is a base-sugar combination. The base portion of the nucleoside is normally a heterocyclic base. The two most common classes of such heterocyclic bases are the purines and the pyrimidines. Nucleotides are nucleosides that further include a phosphate group covalently linked to the sugar portion of the nucleoside. For those nucleosides that include a pentofuranosyl sugar, the phosphate group can be linked to the 2', the 3', or the 5' hydroxyl moiety of the sugar. In forming oligonucleotides, the phosphate groups covalently link adjacent nucleosides to one another to form a linear polymeric compound. In turn, the respective ends of this linear polymeric compound can be further joined to form a circular compound, however, linear compounds are suitable. In addition, linear compounds may have internal nucleotide base complementarity and may therefore fold in a manner as to produce a fully or partially double-stranded compound. Within oligonucleotides, the phosphate groups are commonly referred to as forming the internucleoside backbone of the oligonucleotide. The normal linkage or backbone of RNA and DNA is a 3' to 5' phosphodiester linkage.

[0151] Suitable nucleic acid modifications include, but are not limited to: 2'Omethyl modified nucleotides, 2' Fluoro modified nucleotides, locked nucleic acid (LNA) modified nucleotides, peptide nucleic acid (PNA) modified nucleotides, nucleotides with phosphorothioate linkages,

and a 5' cap (e.g., a 7-methylguanylate cap (m7G)). Additional details and additional modifications are described below.

[0152] A 2'-0-Methyl modified nucleotide (also referred to as 2'-0-Methyl RNA) is a naturally occurring modification of RNA found in tRNA and other small RNAs that arises as a post-transcriptional modification. Oligonucleotides can be directly synthesized that contain 2'-0-Methyl RNA. This modification increases Tm of RNA:RNA duplexes but results in only small changes in RNA:DNA stability. It is stabile with respect to attack by single-stranded ribonucleases and is typically 5 to 10-fold less susceptible to DNases than DNA. It is commonly used in antisense oligos as a means to increase stability and binding affinity to the target message.

[0153] 2' Fluoro modified nucleotides (e.g., 2' Fluoro bases) have a fluorine modified ribose which increases binding affinity (Tm) and also confers some relative nuclease resistance when compared to native RNA. These modifications are commonly employed in ribozymes and siRNAs to improve stability in serum or other biological fluids.

[0154] LNA bases have a modification to the ribose backbone that locks the base in the C3'-endo position, which favors RNA A-type helix duplex geometry. This modification significantly increases Tm and is also very nuclease resistant. Multiple LNA insertions can be placed in an oligo at any position except the 3'-end. Applications have been described ranging from antisense oligos to hybridization probes to SNP detection and allele specific PCR. Due to the large increase in Tm conferred by LNAs, they also can cause an increase in primer dimer formation as well as self -hairpin formation. In some cases, the number of LNAs incorporated into a single oligo is 10 bases or less.

[0155] The phosphorothioate (PS) bond (i.e., a phosphorothioate linkage) substitutes a sulfur atom for a non-bridging oxygen in the phosphate backbone of a nucleic acid (e.g., an oligo). This modification renders the internucleotide linkage resistant to nuclease degradation.

[0156] Phosphorothioate bonds can be introduced between the last 3-5 nucleotides at the 5'- or 3'-end of the oligo to inhibit exonuclease degradation. Including phosphorothioate bonds within the oligo (e.g., throughout the entire oligo) can help reduce attack by endonucleases as well.

[0157] In some embodiments, a subject nucleic acid has one or more nucleotides that are 2'-0-Methyl modified nucleotides. In some embodiments, a subject nucleic acid (e.g., a dsRNA, a siNA, etc.) has one or more 2' Fluoro modified nucleotides. In some embodiments, a subject nucleic acid (e.g., a dsRNA, a siNA, etc.) has one or more LNA bases. In some embodiments, a subject nucleic acid (e.g., a dsRNA, a siNA, etc.) has one or more nucleotides that are linked by a phosphorothioate bond (i.e., the subject nucleic acid has one or more phosphorothioate linkages). In some embodiments, a subject nucleic acid (e.g., a dsRNA, a siNA, etc.) has a 5' cap (e.g., a 7-

methylguanylate cap (m7G)). In some embodiments, a subject nucleic acid (e.g., a dsRNA, a siNA, etc.) has a combination of modified nucleotides. For example, a subject nucleic acid (e.g., a dsRNA, a siNA, etc.) can have a 5' cap (e.g., a 7- methylguanylate cap (m7G)) in addition to having one or more nucleotides with other modifications (e.g., a 2'-0-Methyl nucleotide and/or a 2' Fluoro modified nucleotide and/or a LNA base and/or a phosphorothioate linkage).

[0158] Examples of suitable nucleic acids (e.g., a CasP6 guide RNA) containing modifications include nucleic acids containing modified backbones or non-natural internucleoside linkages. Nucleic acids having modified backbones include those that retain a phosphorus atom in the backbone and those that do not have a phosphorus atom in the backbone.

[0159] Suitable modified oligonucleotide backbones containing a phosphorus atom therein include, for example, phosphorothioates, chiral phosphorothioates, phosphorodithioates, phosphotriesters, aminoalkylphosphotriesters, methyl and other alkyl phosphonates including 3'-alkylene phosphonates, 5'- alkylene phosphonates and chiral phosphonates, phosphinates, phosphoramidates including 3'-amino phosphoramidate and aminoalkylphosphoramidates, phosphorodiamidates, thionophosphoramidates, thionoalkylphosphonates, thionoalkylphosphotriesters, selenophosphates and boranophosphates having normal 3'-5' linkages, 2'-5' linked analogs of these, and those having inverted polarity wherein one or more internucleotide linkages is a 3' to 3', 5' to 5' or 2' to 2' linkage. Suitable oligonucleotides having inverted polarity comprise a single 3' to 3' linkage at the 3'-most internucleotide linkage i.e. a single inverted nucleoside residue which may be a basic (the nucleobase is missing or has a hydroxyl group in place thereof). Various salts (such as, for example, potassium or sodium), mixed salts and free acid forms are also included.

[0160] In some embodiments, a subject nucleic acid comprises one or more phosphorothioate and/or heteroatom internucleoside linkages, in particular $-CH_2-NH-0-CH_2-$, $-CH_2-N(CH_3)-0-CH_2-$ (known as a methylene (methylimino) or MMI backbone), $-CH_2-0-N(CH_3)-CH_2-$, $-CH_2-N(CH_3)-N(CH_3)-CH_2-$ and $-0-N(CH_3)-CH_2-CH_2-$ (wherein the native phosphodiester internucleotide linkage is represented as $-O-P(=0)(OH)-0-CH_2-$). MMI type internucleoside linkages are disclosed in the above referenced U.S. Pat. No. 5,489,677, the disclosure of which is incorporated herein by reference in its entirety. Suitable amide internucleoside linkages are disclosed in U.S. Pat. No. 5,602,240, the disclosure of which is incorporated herein by reference in its entirety.

[0161] Also suitable are nucleic acids having morpholino backbone structures as described in, e.g., U.S. Pat. No. 5,034,506. For example, in some embodiments, a subject nucleic acid comprises a 6- membered morpholino ring in place of a ribose ring. In some of these

embodiments, a phosphorodiamidate or other non-phosphodiester internucleoside linkage replaces a phosphodiester linkage.

[0162] Suitable modified polynucleotide backbones that do not include a phosphorus atom therein have backbones that are formed by short chain alkyl or cycloalkyl internucleoside linkages, mixed heteroatom and alkyl or cycloalkyl internucleoside linkages, or one or more short chain heteroatomic or heterocyclic internucleoside linkages. These include those having morpholino linkages (formed in part from the sugar portion of a nucleoside); siloxane backbones; sulfide, sulfoxide and sulfone backbones; formacetyl and thioformacetyl backbones; methylene formacetyl and thioformacetyl backbones; riboacetyl backbones; alkene containing backbones; sulfamate backbones; methyleneimino and methylenehydrazino backbones; sulfonate and sulfonamide backbones; amide backbones; and others having mixed N, O, S and CH_2 component parts.

[0163] A subject nucleic acid can be a nucleic acid mimetic. The term "mimetic" as it is applied to polynucleotides is intended to include polynucleotides wherein only the furanose ring or both the furanose ring and the internucleotide linkage are replaced with non-fur anose groups, replacement of only the furanose ring is also referred to in the art as being a sugar surrogate. The heterocyclic base moiety or a modified heterocyclic base moiety is maintained for hybridization with an appropriate target nucleic acid. One such nucleic acid, a polynucleotide mimetic that has been shown to have excellent hybridization properties, is referred to as a peptide nucleic acid (PNA). In PNA, the sugar-backbone of a polynucleotide is replaced with an amide containing backbone, in particular an aminoethylglycine backbone. The nucleotides are retained and are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone.

[0164] One polynucleotide mimetic that has been reported to have excellent hybridization properties is a peptide nucleic acid (PNA). The backbone in PNA compounds is two or more linked aminoethylglycine units which gives PNA an amide containing backbone. The heterocyclic base moieties are bound directly or indirectly to aza nitrogen atoms of the amide portion of the backbone. Representative U.S. patents that describe the preparation of PNA compounds include, but are not limited to: U.S. Pat. Nos. 5,539,082; 5,714,331 ; and 5,719,262, the disclosures of which are incorporated herein by reference in their entirety.

[0165] Another class of polynucleotide mimetic that has been studied is based on linked morpholino units (morpholino nucleic acid) having heterocyclic bases attached to the morpholino ring. A number of linking groups have been reported that link the morpholino monomeric units in a morpholino nucleic acid. One class of linking groups has been selected to give a non-ionic oligomeric compound. The non- ionic morpholino-based oligomeric compounds are less likely to

have undesired interactions with cellular proteins. Morpholino-based polynucleotides are non-ionic mimics of oligonucleotides which are less likely to form undesired interactions with cellular proteins (Dwaine A. Braasch and David R. Corey, Biochemistry, 2002, 41(14), 4503-4510). Morpholino-based polynucleotides are disclosed in U.S. Pat. No. 5,034,506, the disclosure of which is incorporated herein by reference in its entirety. A variety of compounds within the morpholino class of polynucleotides have been prepared, having a variety of different linking groups joining the monomeric subunits.

[0166] A further class of polynucleotide mimetic is referred to as cyclohexenyl nucleic acids (CeNA). The furanose ring normally present in a DNA/RNA molecule is replaced with a cyclohexenyl ring. CeNA DMT protected phosphoramidite monomers have been prepared and used for oligomeric compound synthesis following classical phosphoramidite chemistry. Fully modified CeNA oligomeric compounds and oligonucleotides having specific positions modified with CeNA have been prepared and studied (see Wang et al., /. Am. Chem. Soc, 2000, 122, 8595-8602, the disclosure of which is incorporated herein by reference in its entirety). In general the incorporation of CeNA monomers into a DNA chain increases its stability of a DNA/RNA hybrid. CeNA oligoadenylates formed complexes with RNA and DNA complements with similar stability to the native complexes. The study of incorporating CeNA structures into natural nucleic acid structures was shown by NMR and circular dichroism to proceed with easy conformational adaptation.

[0167] A further modification includes Locked Nucleic Acids (LNAs) in which the 2'-hydroxyl group is linked to the 4' carbon atom of the sugar ring thereby forming a 2'-C,4'-C-oxymethylene linkage thereby forming a bicyclic sugar moiety. The linkage can be a methylene (-CH_2-), group bridging the 2' oxygen atom and the 4' carbon atom wherein n is 1 or 2 (Singh et al., Chem. Commun., 1998, 4, 455-456, the disclosure of which is incorporated herein by reference in its entirety). LNA and LNA analogs display very high duplex thermal stabilities with complementary DNA and RNA (Tm=+3 to +10° C), stability towards 3'-exonucleolytic degradation and good solubility properties. Potent and nontoxic antisense oligonucleotides containing LNAs have been described (e.g., Wahlestedt et al., Proc. Natl. Acad. Sci. U.S.A., 2000, 97, 5633-5638, the disclosure of which is incorporated herein by reference in its entirety).

[0168] The synthesis and preparation of the LNA monomers adenine, cytosine, guanine, 5-methyl- cytosine, thymine and uracil, along with their oligomerization, and nucleic acid recognition properties have been described (e.g., Koshkin et al., Tetrahedron, 1998, 54, 3607-3630, the disclosure of which is incorporated herein by reference in its entirety). LNAs and preparation thereof are also described in WO 98/39352 and WO 99/14226, as well as U.S.

applications 20120165514, 20100216983, 20090041809, 20060117410, 20040014959, 20020094555, and 20020086998, the disclosures of which are incorporated herein by reference in their entirety.

[0169] A subject nucleic acid can also include one or more substituted sugar moieties. Suitable polynucleotides comprise a sugar substituent group selected from: OH; F; O-, S-, or N-alkyl; O-, S-, or N- alkenyl; O-, S- or N-alkynyl; or O-alkyl-O-alkyl, wherein the alkyl, alkenyl and alkynyl may be substituted or unsubstituted C.sub. l to do alkyl or C_2 to do alkenyl and alkynyl. Particularly suitable are $O((CH_2)_nO)_mCH_3$, $O(CH_2)_nOCH_3$, $O(CH_2)_nNH_2$, $O(CH_2)_nCH_3$, $O(CH_2)_nONH_2$, and $O(CH_2)_nON((CH_2)_nCH_3)_2$, where n and m are from 1 to about 10. Other suitable polynucleotides comprise a sugar substituent group selected from: d to C_{10} lower alkyl, substituted lower alkyl, alkenyl, alkynyl, alkaryl, aralkyl, O-alkaryl or O-aralkyl, SH, SCH_3, OCN, CI, Br, CN, CF_3, OCF_3, $SOCH_3$, SO_2CH_3, ONO_2, NO_2, N_3, NH_2, heterocycloalkyl, heterocycloalkaryl, aminoalkylamino, polyalkylamino, substituted silyl, an RNA cleaving group, a reporter group, an intercalator, a group for improving the pharmacokinetic properties of an oligonucleotide, or a group for improving the pharmacodynamic properties of an oligonucleotide, and other substituents having similar properties. A suitable modification includes 2'-methoxyethoxy (2'-O-CH_2 CH_2OCH_3, also known as 2'-O-(2-methoxyethyl) or 2'-MOE) (Martin et al., Helv. Chim. Acta, 1995, 78, 486-504, the disclosure of which is incorporated herein by reference in its entirety) i.e., an alkoxyalkoxy group. A further suitable modification includes 2'-dimethylaminooxyethoxy, i.e., a $O(CH_2)_2ON(CH_3)_2$ group, also known as 2'-DMAOE, as described in examples herein below, and 2'-dimethylaminoethoxyethoxy (also known in the art as 2'-O-dimethyl- amino-ethoxy-ethyl or 2'-DMAEOE), i.e., 2'-O-CH_2-O-CH_2-$N(CH_3)_2$.

[0170] Other suitable sugar substituent groups include methoxy (-O-CH_3), aminopropoxy (—O CH_2 CH_2 CH_2NH_2), allyl (-CH_2-CH=CH_2), -O-allyl (-O-- CH_2— CH=CH_2) and fluoro (F). 2'-sugar substituent groups may be in the arabino (up) position or ribo (down) position. A suitable 2'-arabino modification is 2'-F. Similar modifications may also be made at other positions on the oligomeric compound, particularly the 3' position of the sugar on the 3' terminal nucleoside or in 2'-5' linked oligonucleotides and the 5' position of 5' terminal nucleotide. Oligomeric compounds may also have sugar mime tics such as cyclobutyl moieties in place of the pentofuranosyl sugar.

[0171] A subject nucleic acid may also include nucleobase (often referred to in the art simply as "base") modifications or substitutions. As used herein, "unmodified" or "natural" nucleobases include the purine bases adenine (A) and guanine (G), and the pyrimidine bases thymine (T), cytosine (C) and uracil (U). Modified nucleobases include other synthetic and natural nucleobases such as 5-methylcytosine (5- me-C), 5-hydroxymethyl cytosine, xanthine,

hypoxanthine, 2-aminoadenine, 6-methyl and other alkyl derivatives of adenine and guanine, 2-propyl and other alkyl derivatives of adenine and guanine, 2- thiouracil, 2-thiothymine and 2-thiocytosine, 5-halouracil and cytosine, 5-propynyl (-C=C-CH_3) uracil and cytosine and other alkynyl derivatives of pyrimidine bases, 6-azo uracil, cytosine and thymine, 5- uracil (pseudouracil), 4-thiouracil, 8-halo, 8-amino, 8-thiol, 8-thioalkyl, 8-hydroxyl and other 8-substituted adenines and guanines, 5-halo particularly 5-bromo, 5-trifluoromethyl and other 5-substituted uracils and cytosines, 7-methylguanine and 7-methyladenine, 2-F-adenine, 2-amino-adenine, 8- azaguanine and 8-azaadenine, 7-deazaguanine and 7-deazaadenine and 3-deazaguanine and 3- deazaadenine. Further modified nucleobases include tricyclic pyrimidines such as phenoxazine cytidine(lH-pyrirnido(5,4-b)(l,4)benzoxazin-2(3H)-one), phenothiazine cytidine (lH-pyrimido(5,4- b)(l,4)benzothiazin-2(3H)-one), G-clamps such as a substituted phenoxazine cytidine (e.g. 9-(2- aminoethoxy)-H-pyrimido(5,4-(b) (l,4)benzoxazin-2(3H)-one), carbazole cytidine (2H-pyrimido(4,5- b)indol-2-one), pyridoindole cytidine (H-pyrido(3',2':4,5)pyrrolo(2,3-d)pyrimidin-2-one).

[0172] Heterocyclic base moieties may also include those in which the purine or pyrimidine base is replaced with other heterocycles, for example 7-deaza-adenine, 7-deazaguanosine, 2-aminopyridine and 2-pyridone. Further nucleobases include those disclosed in U.S. Pat. No. 3,687,808, those disclosed in The Concise Encyclopedia Of Polymer Science And Engineering, pages 858-859, Kroschwitz, J. I., ed. John Wiley & Sons, 1990, those disclosed by Englisch et al., Angewandte Chemie, International Edition, 1991, 30, 613, and those disclosed by Sanghvi, Y. S., Chapter 15, Antisense Research and Applications, pages 289-302, Crooke, S. T. and Lebleu, B., ed., CRC Press, 1993; the disclosures of which are incorporated herein by reference in their entirety. Certain of these nucleobases are useful for increasing the binding affinity of an oligomeric compound. These include 5-substituted pyrimidines, 6- azapyrimidines and N-2, N-6 and O-6 substituted purines, including 2-aminopropyladenine, 5- propynyluracil and 5-propynylcytosine. 5-methylcytosine substitutions have been shown to increase nucleic acid duplex stability by 0.6-1.2° C. (Sanghvi et al., eds., Antisense Research and Applications, CRC Press, Boca Raton, 1993, pp. 276-278; the disclosure of which is incorporated herein by reference in its entirety) and are suitable base substitutions, e.g., when combined with 2'-O-methoxyethyl sugar modifications.

[0173] Another possible modification of a subject nucleic acid involves chemically linking to the polynucleotide one or more moieties or conjugates which enhance the activity, cellular distribution or cellular uptake of the oligonucleotide. These moieties or conjugates can include conjugate groups covalently bound to functional groups such as primary or secondary hydroxyl

groups. Conjugate groups include, but are not limited to, intercalators, reporter molecules, polyamines, polyamides, polyethylene glycols, polyethers, groups that enhance the pharmacodynamic properties of oligomers, and groups that enhance the pharmacokinetic properties of oligomers. Suitable conjugate groups include, but are not limited to, cholesterols, lipids, phospholipids, biotin, phenazine, folate, phenanthridine, anthraquinone, acridine, fluoresceins, rhodamines, coumarins, and dyes. Groups that enhance the pharmacodynamic properties include groups that improve uptake, enhance resistance to degradation, and/or strengthen sequence-specific hybridization with the target nucleic acid. Groups that enhance the pharmacokinetic properties include groups that improve uptake, distribution, metabolism or excretion of a subject nucleic acid.

[0174] Conjugate moieties include but are not limited to lipid moieties such as a cholesterol moiety (Letsinger et al., Proc. Natl. Acad. Sci. USA, 1989, 86, 6553-6556), cholic acid (Manoharan et al., Bioorg. Med. Chem. Let., 1994, 4, 1053-1060), a thioether, e.g., hexyl-S-tritylthiol (Manoharan et al., Ann. N Y. Acad. Sci., 1992, 660, 306-309; Manoharan et al., Bioorg. Med. Chem. Let., 1993, 3, 2765-2770), a thiocholesterol (Oberhauser et al., Nucl. Acids Res., 1992, 20, 533-538), an aliphatic chain, e.g., dodecandiol or undecyl residues (Saison-Behmoaras et al., EMBO J., 1991, 10, 1111-1118; Kabanov et al, FEBS Lett., 1990, 259, 327-330; Svinarchuk et al., Biochimie, 1993, 75, 49-54), a phospholipid, e.g., di-hexadecyl-rac-glycerol or triethylammonium 1,2-di-0-hexadecyl-rac-glycero-3-H-phosphonate (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654; Shea et al., Nucl. Acids Res., 1990, 18, 3777- 3783), a polyamine or a polyethylene glycol chain (Manoharan et al., Nucleosides & Nucleotides, 1995, 14, 969-973), or adamantane acetic acid (Manoharan et al., Tetrahedron Lett., 1995, 36, 3651-3654), a palmityl moiety (Mishra et al., Biochim. Biophys. Acta, 1995, 1264, 229-237), or an octadecylamine or hexylamino-carbonyl-oxycholesterol moiety (Crooke et al., /. Pharmacol. Exp. Ther., 1996, 277, 923- 937).

[0175] A conjugate may include a "Protein Transduction Domain" or PTD (also known as a CPP - cell penetrating peptide), which may refer to a polypeptide, polynucleotide, carbohydrate, or organic or inorganic compound that facilitates traversing a lipid bilayer, micelle, cell membrane, organelle membrane, or vesicle membrane. A PTD attached to another molecule, which can range from a small polar molecule to a large macromolecule and/or a nanoparticle, facilitates the molecule traversing a membrane, for example going from extracellular space to intracellular space, or cytosol to within an organelle (e.g., the nucleus). In some embodiments, a PTD is covalently linked to the 3' end of an exogenous polynucleotide. In some embodiments, a PTD is covalently linked to the 5' end of an exogenous polynucleotide. Exemplary PTDs include but are

not limited to a minimal undecapeptide protein transduction domain (corresponding to residues 47-57 of HIV- 1 TAT comprising SEQ ID NO: 49); a polyarginine sequence comprising a number of arginines sufficient to direct entry into a cell (e.g., 3, 4, 5, 6, 7, 8, 9, 10, or 10-50 arginines); a VP22 domain (Zender et al. (2002) Cancer Gene Ther. 9(6):489-96); an Drosophila Antennapedia protein transduction domain (Noguchi et al. (2003) Diabetes 52(7): 1732-1737); a truncated human calcitonin peptide (Trehin et al. (2004) Pharm. Research 21 : 1248-1256); polylysine (Wender et al. (2000) Proc. Natl. Acad. Sci. USA 97: 13003-13008); SEQ ID NO: 50; Transportan (SEQ ID NO: 51; SEQ ID NO: 52; and SEQ ID NO: 53. Exemplary PTDs include but are not limited to, SEQ ID NO: 54; an arginine homopolymer of from 3 arginine residues to 50 arginine residues. Exemplary PTD domain amino acid sequences include, but are not limited to, any of the following: SEQ ID NO: 55; SEQ ID NO: 56; SEQ ID NO: 57; and SEQ ID NO: 58. In some embodiments, the PTD is an activatable CPP (ACPP) (Aguilera et al. (2009) Integr Biol (Camb) June; 1(5-6): 371-381). ACPPs comprise a polycationic CPP (e.g., Arg9 or "R9") connected via a cleavable linker to a matching polyanion (e.g., Glu9 or "E9"), which reduces the net charge to nearly zero and thereby inhibits adhesion and uptake into cells. Upon cleavage of the linker, the polyanion is released, locally unmasking the polyarginine and its inherent adhesiveness, thus "activating" the ACPP to traverse the membrane.

[0176] A CasP6 guide RNA (or a nucleic acid comprising a nucleotide sequence encoding same) and/or a CasP6 polypeptide of the present disclosure (or a nucleic acid comprising a nucleotide sequence encoding same) and/or a CasP6 fusion polypeptide of the present disclosure (or a nucleic acid that includes a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure) and/or a donor polynucleotide (donor template) can be introduced into a host cell by any of a variety of well-known methods.

[0177] Any of a variety of compounds and methods can be used to deliver to a target cell a CasP6 system of the present disclosure (e.g., where a CasP6 system comprises: a) a CasP6 polypeptide of the present disclosure and a CasP6 guide RNA; b) a CasP6 polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; c) a CasP6 fusion polypeptide of the present disclosure and a CasP6 guide RNA; d) a CasP6 fusion polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; e) an mRNA encoding a CasP6 polypeptide of the present disclosure; and a CasP6 guide RNA; f) an mRNA encoding a CasP6 polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; g) an mRNA encoding a CasP6 fusion polypeptide of the present disclosure; and a CasP6 guide RNA; h) an mRNA encoding a CasP6 fusion polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; i) a recombinant

expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure and a nucleotide sequence encoding a CasP6 guide RNA; j) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, a nucleotide sequence encoding a CasP6 guide RNA, and a nucleotide sequence encoding a donor template nucleic acid; k) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure and a nucleotide sequence encoding a CasP6 guide RNA; 1) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, a nucleotide sequence encoding a CasP6 guide RNA, and a nucleotide sequence encoding a donor template nucleic acid; m) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; n) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; and a donor template nucleic acid; o) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; p) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; and a donor template nucleic acid; q) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, a nucleotide sequence encoding a first CasP6 guide RNA, and a nucleotide sequence encoding a second CasP6 guide RNA; or r) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, a nucleotide sequence encoding a first CasP6 guide RNA, and a nucleotide sequence encoding a second CasP6 guide RNA; or some variation of one of (a) through (r). As a non-limiting example, a CasP6 system of the present disclosure can be combined with a lipid. As another non-limiting example, a CasP6 system of the present disclosure can be combined with a particle, or formulated into a particle.

[0178] Methods of introducing a nucleic acid into a host cell are known in the art, and any convenient method can be used to introduce a subject nucleic acid (e.g., an expression construct/vector) into a target cell (e.g., prokaryotic cell, eukaryotic cell, plant cell, animal cell, mammalian cell, human cell, and the like). Suitable methods include, e.g., viral infection, transfection, conjugation, protoplast fusion, lipofection, electroporation, calcium phosphate

precipitation, polyethyleneimine (PEI) -mediated transfection, DEAE-dextran mediated transfection, liposome-mediated transfection, particle gun technology, calcium phosphate precipitation, direct micro injection, nanoparticle-mediated nucleic acid delivery (see, e.g., Panyam et., al Adv Drug Deliv Rev. 2012 Sep 13. pii: S0169-409X(12)00283-9. doi: 10.1016/j.addr.2012.09.023), and the like.

[0179] In some cases, a CasP6 polypeptide of the present disclosure is provided as a nucleic acid (e.g., an mRNA, a DNA, a plasmid, an expression vector, a viral vector, etc.) that encodes the CasP6 polypeptide. In some cases, the CasP6 polypeptide of the present disclosure is provided directly as a protein (e.g., without an associated guide RNA or with an associate guide RNA, i.e., as a ribonucleoprotein complex). A CasP6 polypeptide of the present disclosure can be introduced into a cell (provided to the cell) by any convenient method; such methods are known to those of ordinary skill in the art. As an illustrative example, a CasP6 polypeptide of the present disclosure can be injected directly into a cell (e.g., with or without a CasP6 guide RNA or nucleic acid encoding a CasP6 guide RNA, and with or without a donor polynucleotide). As another example, a preformed complex of a CasP6 polypeptide of the present disclosure and a CasP6 guide RNA (an RNP) can be introduced into a cell (e.g, eukaryotic cell) (e.g., via injection, via nucleofection; via a protein transduction domain (PTD) conjugated to one or more components, e.g., conjugated to the CasP6 protein, conjugated to a guide RNA, conjugated to a CasP6 polypeptide of the present disclosure and a guide RNA; etc.).

[0180] In some cases, a CasP6 fusion polypeptide (e.g., dCasP6 fused to a fusion partner, nickase CasP6 fused to a fusion partner, etc.) of the present disclosure is provided as a nucleic acid (e.g., an mRNA, a DNA, a plasmid, an expression vector, a viral vector, etc.) that encodes the CasP6 fusion polypeptide. In some cases, the CasP6 fusion polypeptide of the present disclosure is provided directly as a protein (e.g., without an associated guide RNA or with an associate guide RNA, i.e., as a ribonucleoprotein complex). A CasP6 fusion polypeptide of the present disclosure can be introduced into a cell (provided to the cell) by any convenient method; such methods are known to those of ordinary skill in the art. As an illustrative example, a CasP6 fusion polypeptide of the present disclosure can be injected directly into a cell (e.g., with or without nucleic acid encoding a CasP6 guide RNA and with or without a donor polynucleotide). As another example, a preformed complex of a CasP6 fusion polypeptide of the present disclosure and a CasP6 guide RNA (an RNP) can be introduced into a cell (e.g., via injection, via nucleofection; via a protein transduction domain (PTD) conjugated to one or more components, e.g., conjugated to the CasP6 fusion protein, conjugated to a guide RNA, conjugated to a CasP6 fusion polypeptide of the present disclosure and a guide RNA; etc.).

[0181] In some cases, a nucleic acid (e.g., a CasP6 guide RNA; a nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure; etc.) is delivered to a cell (e.g., a target host cell) and/or a polypeptide (e.g., a CasP6 polypeptide; a CasP6 fusion polypeptide) in a particle, or associated with a particle. In some cases, a CasP6 system of the present disclosure is delivered to a cell in a particle, or associated with a particle. The terms "particle" and nanoparticle" can be used interchangeable, as appropriate. A recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure and/or a CasP6 guide RNA, an mRNA comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, and guide RNA may be delivered simultaneously using particles or lipid envelopes; for instance, a CasP6 polypeptide and a CasP6 guide RNA, e.g., as a complex (e.g., a ribonucleoprotein (RNP) complex), can be delivered via a particle, e.g., a delivery particle comprising lipid or lipidoid and hydrophilic polymer, e.g., a cationic lipid and a hydrophilic polymer, for instance wherein the cationic lipid comprises l,2-dioleoyl-3-trimethylammonium-propane (DOTAP) or l,2-ditetradecanoyl-sn-glycero-3-phosphocholine (DMPC) and/or wherein the hydrophilic polymer comprises ethylene glycol or polyethylene glycol (PEG); and/or wherein the particle further comprises cholesterol (e.g., particle from formulation l=DOTAP 100, DMPC 0, PEG 0, Cholesterol 0; formulation number 2=DOTAP 90, DMPC 0, PEG 10, Cholesterol 0; formulation number 3=DOTAP 90, DMPC 0, PEG 5, Cholesterol 5). For example, a particle can be formed using a multistep process in which a CasP6 polypeptide and a CasP6 guideRNA are mixed together, e.g., at a 1: 1 molar ratio, e.g., at room temperature, e.g., for 30 minutes, e.g., in sterile, nuclease free 1 x phosphate-buffered saline (PBS); and separately, DOTAP, DMPC, PEG, and cholesterol as applicable for the formulation are dissolved in alcohol, e.g., 100% ethanol; and, the two solutions are mixed together to form particles containing the complexes).

[0182] A CasP6 polypeptide of the present disclosure (or an mRNA comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure; or a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure) and/or CasP6 guide RNA (or a nucleic acid such as one or more expression vectors encoding the CasP6 guide RNA) may be delivered simultaneously using particles or lipid envelopes. For example, a biodegradable core- shell structured nanoparticle with a poly (β-amino ester) (PBAE) core enveloped by a phospholipid bilayer shell can be used. In some cases, particles/nanoparticles based on self assembling bioadhesive polymers are used; such particles/nanoparticles may be applied to oral delivery of peptides, intravenous delivery of peptides and nasal delivery of peptides, e.g., to the brain. Other embodiments, such as oral absorption and ocular delivery of

hydrophobic drugs are also contemplated. A molecular envelope technology, which involves an engineered polymer envelope which is protected and delivered to the site of the disease, can be used. Doses of about 5 mg/kg can be used, with single or multiple doses, depending on various factors, e.g., the target tissue.

[0183] Lipidoid compounds (e.g., as described in US patent application 20110293703) are also useful in the administration of polynucleotides, and can be used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure.

[0184] In one aspect, the aminoalcohol lipidoid compounds are combined with an agent to be delivered to a cell or a subject to form microparticles, nanoparticles, liposomes, or micelles. The aminoalcohol lipidoid compounds may be combined with other aminoalcohol lipidoid compounds, polymers (synthetic or natural), surfactants, cholesterol, carbohydrates, proteins, lipids, etc. to form the particles. These particles may then optionally be combined with a pharmaceutical excipient to form a pharmaceutical composition.

[0185] A poly(beta-amino alcohol) (PBAA) can be used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. US Patent Publication No. 20130302401 relates to a class of poly(beta-amino alcohols) (PBAAs) that has been prepared using combinatorial polymerization.

[0186] Sugar-based particles may be used, for example GalNAc, as described with reference to WO2014118272 (incorporated herein by reference) and Nair, J K et al., 2014, Journal of the American Chemical Society 136 (49), 16958-16961) can be used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell.

[0187] In some cases, lipid nanoparticles (LNPs) are used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. Negatively charged polymers such as RNA may be loaded into LNPs at low pH values (e.g., pH 4) where the ionizable lipids display a positive charge. However, at physiological pH values, the LNPs exhibit a low surface charge compatible with longer circulation times. Four species of ionizable cationic lipids have been focused upon, namely 1,2-dilineoyl-3-dimethylammonium-propane (DLinDAP), 1,2- dilinoleyloxy-3-N,N-dimethylaminopropane

(DLinDMA), 1,2-dilinoleyloxy-keto-N,N-dimethyl-3- aminopropane (DLinKDMA), and l,2-dilinoleyl-4-(2-dimethylaminoethyl)-[l,3]-dioxolane (DLinKC2- DMA). Preparation of LNPs and is described in, e.g., Rosin et al. (2011) Molecular Therapy 19: 1286- 2200). The cationic lipids l,2-dilineoyl-3-dimethylammonium-propane (DLinDAP), l,2-dilinoleyloxy-3- N,N-dimethylaminopropane (DLinDMA), l,2-dilinoleyloxyketo-N,N-dimethyl-3-aminopropane (DLinKDMA), l,2-dilinoleyl-4-(2-dimethylaminoethyl)-[l,3]-dioxolane (DLinKC2-DM A), (3-o-[2"- (methoxypolyethyleneglycol 2000) succinoyl]-l,2-dimyristoyl-sn-glycol (PEG-S-DMG), and R-3- [(.omega.-methoxy-poly(ethylene glycol)2000) carbamoyl] -l,2-dimyristyloxlpropyl-3 - amine (PEG-C- DOMG) may be used. A nucleic acid (e.g., a CasP6 guide RNA; a nucleic acid of the present disclosure; etc.) may be encapsulated in LNPs containing DLinDAP, DLinDMA, DLinK-DMA, and DLinKC2-DMA (cationic lipid:DSPC:CHOL: PEGS-DMG or PEG-C-DOMG at 40: 10:40: 10 molar ratios). In some cases, 0.2% SP-DiOC18 is incorporated.

[0188] Spherical Nucleic Acid (SNA™) constructs and other nanoparticles (particularly gold nanoparticles) can be used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell.. See, e.g., Cutler et al., J. Am. Chem. Soc. 2011 133:9254-9257, Hao et al., Small. 2011 7:3158-3162, Zhang et al., ACS Nano. 2011 5:6962-6970, Cutler et al., J. Am. Chem. Soc. 2012 134: 1376-1391, Young et al., Nano Lett. 2012 12:3867-71, Zheng et al., Proc. Natl. Acad. Sci. USA. 2012 109: 11975-80, Mirkin, Nanomedicine 2012 7:635-638 Zhang et al., J. Am. Chem. Soc. 2012 134:16488-1691, Weintraub, Nature 2013 495:S14-S16, Choi et al., Proc. Natl. Acad. Sci. USA. 2013 110(19): 7625-7630, Jensen et al, Sci. Transl. Med. 5, 209ral52 (2013) and Mirkin, et al., Small, 10: 186-192.

[0189] Self-assembling nanoparticles with RNA may be constructed with polyethyleneimine (PEI) that is PEGylated with an Arg-Gly-Asp (RGD) peptide ligand attached at the distal end of the polyethylene glycol (PEG).

[0190] In general, a "nanoparticle" refers to any particle having a diameter of less than 1000 nm. In some cases, nanoparticles suitable for use in delivering a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell have a diameter of 500 nm or less, e.g., from 25 nm to 35 nm, from 35 nm to 50 nm, from 50 nm to 75 nm, from 75 nm to 100 nm, from 100 nm to 150 nm, from 150 nm to 200 nm, from 200 nm to 300 nm, from 300 nm to 400 nm, or from 400 nm to 500 nm. In some cases, nanoparticles suitable for use in delivering a CasP6 polypeptide of the present disclosure, a

CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell have a diameter of from 25 nm to 200 nm. In some cases, nanoparticles suitable for use in delivering a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell have a diameter of 100 nm or less In some cases, nanoparticles suitable for use in delivering a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell have a diameter of from 35 nm to 60 nm.

[0191] Nanoparticles suitable for use in delivering a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell may be provided in different forms, e.g., as solid nanoparticles (e.g., metal such as silver, gold, iron, titanium), non-metal, lipid-based solids, polymers), suspensions of nanoparticles, or combinations thereof. Metal, dielectric, and semiconductor nanoparticles may be prepared, as well as hybrid structures (e.g., core-shell nanoparticles). Nanoparticles made of semiconducting material may also be labeled quantum dots if they are small enough (typically below 10 nm) that quantization of electronic energy levels occurs. Such nanoscale particles are used in biomedical applications as drug carriers or imaging agents and may be adapted for similar purposes in the present disclosure.

[0192] Semi-solid and soft nanoparticles are also suitable for use in delivering a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. A prototype nanoparticle of semi-solid nature is the liposome.

[0193] In some cases, an exosome is used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. Exosomes are endogenous nano-vesicles that transport RNAs and proteins, and which can deliver RNA to the brain and other target organs.

[0194] In some cases, a liposome is used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. Liposomes are spherical vesicle structures composed of a uni- or multilamellar lipid bilayer

surrounding internal aqueous compartments and a relatively impermeable outer lipophilic phospholipid bilayer. Liposomes can be made from several different types of lipids; however, phospholipids are most commonly used to generate liposomes. Although liposome formation is spontaneous when a lipid film is mixed with an aqueous solution, it can also be expedited by applying force in the form of shaking by using a homogenizer, sonicator, or an extrusion apparatus. Several other additives may be added to liposomes in order to modify their structure and properties. For instance, either cholesterol or sphingomyelin may be added to the liposomal mixture in order to help stabilize the liposomal structure and to prevent the leakage of the liposomal inner cargo. A liposome formulation may be mainly comprised of natural phospholipids and lipids such as 1,2-distearoryl-sn-glycero-3 -phosphatidyl choline (DSPC), sphingomyelin, egg phosphatidylcholines and monosialoganglioside.

[0195] A stable nucleic-acid-lipid particle (SNALP) can be used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. The SNALP formulation may contain the lipids 3-N-[(methoxypoly(ethylene glycol) 2000) carbamoyl] -1,2-dimyristyloxy-propylamine (PEG-C-DMA), l,2-dilinoleyloxy-N,N-dimethyl-3- aminopropane (DLinDMA), l,2-distearoyl-sn-glycero-3-phosphocholine (DSPC) and cholesterol, in a 2:40: 10:48 molar percent ratio. The SNALP liposomes may be prepared by formulating D-Lin-DMA and PEG-C-DMA with distearoylphosphatidylcholine (DSPC), Cholesterol and siRNA using a 25: 1 lipid/siRNA ratio and a 48/40/10/2 molar ratio of Cholesterol/D-Lin-DMA/DSPC/PEG-C-DMA. The resulting SNALP liposomes can be about 80-100 nm in size. A SNALP may comprise synthetic cholesterol (Sigma-Aldrich, St Louis, Mo., USA), dipalmitoylphosphatidylcholine (Avanti Polar Lipids, Alabaster, Ala., USA), 3-N-[(w-methoxy poly(ethylene glycol)2000)carbamoyl]-l,2-dimyrestyloxypropylamine, and cationic l,2-dilinoleyloxy-3-N,Ndimethylaminopropane. A SNALP may comprise synthetic cholesterol (Sigma-Aldrich), l,2-distearoyl-sn-glycero-3-phosphocholine (DSPC; Avanti Polar Lipids Inc.), PEG-cDMA, and l,2-dilinoleyloxy-3-(N;N-dimethyl)aminopropane (DLinDMA).

[0196] Other cationic lipids, such as amino lipid 2,2-dilinoleyl-4-dimethylaminoethyl-[l,3]-dioxolane (DLin-KC2-DMA) can be used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. A preformed vesicle with the following lipid composition may be contemplated: amino lipid, distearoylphosphatidylcholine (DSPC), cholesterol and (R)-2,3-bis(octadecyloxy)

propyl- l-(methoxy poly(ethyleneglycol)2000)propylcarbamate (PEG-lipid) in the molar ratio 40/10/40/10, respectively, and a FVII siRN A/total lipid ratio of approximately 0.05 (w/w). To ensure a narrow particle size distribution in the range of 70-90 nm and a low polydispersity index of 0.11.+-.0.04 (n=56), the particles may be extruded up to three times through 80 nm membranes prior to adding the guide RNA. Particles containing the highly potent amino lipid 16 may be used, in which the molar ratio of the four lipid components 16, DSPC, cholesterol and PEG-lipid (50/10/38.5/1.5) which may be further optimized to enhance in vivo activity.

[0197] Lipids may be formulated with a CasP6 system of the present disclosure or component(s) thereof or nucleic acids encoding the same to form lipid nanoparticles (LNPs). Suitable lipids include, but are not limited to, DLin-KC2-DMA4, CI 2-200 and colipids disteroylphosphatidyl choline, cholesterol, and PEG-DMG may be formulated with a CasP6 system, or component thereof, of the present disclosure, using a spontaneous vesicle formation procedure. The component molar ratio may be about 50/10/38.5/1.5 (DLin-KC2-DMA or C12-200/disteroylphosphatidyl choline/cholesterol/PEG-DMG).

[0198] A CasP6 system of the present disclosure, or a component thereof, may be delivered encapsulated in PLGA microspheres such as that further described in US published applications 20130252281 and 20130245107 and 20130244279.

[0199] Supercharged proteins can be used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. Supercharged proteins are a class of engineered or naturally occurring proteins with unusually high positive or negative net theoretical charge. Both supernegatively and superpositively charged proteins exhibit the ability to withstand thermally or chemically induced aggregation. Superpositively charged proteins are also able to penetrate mammalian cells. Associating cargo with these proteins, such as plasmid DNA, RNA, or other proteins, can facilitate the functional delivery of these macromolecules into mammalian cells both in vitro and in vivo.

[0200] Cell Penetrating Peptides (CPPs) can be used to deliver a CasP6 polypeptide of the present disclosure, a CasP6 fusion polypeptide of the present disclosure, an RNP of the present disclosure, a nucleic acid of the present disclosure, or a CasP6 system of the present disclosure, to a target cell. CPPs typically have an amino acid composition that either contains a high relative abundance of positively charged amino acids such as lysine or arginine or has sequences that contain an alternating pattern of polar/charged amino acids and non-polar, hydrophobic amino acids.

[0201] The present disclosure provides a modified cell comprising a CasP6 polypeptide of the present disclosure and/or a nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure. The present disclosure provides a modified cell comprising a CasP6 polypeptide of the present disclosure, where the modified cell is a cell that does not normally comprise a CasP6 polypeptide of the present disclosure. The present disclosure provides a modified cell (e.g., a genetically modified cell) comprising nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure. The present disclosure provides a genetically modified cell that is genetically modified with an mRNA comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure. The present disclosure provides a genetically modified cell that is genetically modified with a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure. The present disclosure provides a genetically modified cell that is genetically modified with a recombinant expression vector comprising: a) a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure; and b) a nucleotide sequence encoding a CasP6 guide RNA of the present disclosure. The present disclosure provides a genetically modified cell that is genetically modified with a recombinant expression vector comprising: a) a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure; b) a nucleotide sequence encoding a CasP6 guide RNA of the present disclosure; and c) a nucleotide sequence encoding a donor template.

[0202] A cell that serves as a recipient for a CasP6 polypeptide of the present disclosure and/or a nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure and/or a CasP6 guide RNA of the present disclosure, can be any of a variety of cells, including, e.g., in vitro cells; in vivo cells; ex vivo cells; primary cells; cancer cells; animal cells; plant cells; algal cells; fungal cells; etc. A cell that serves as a recipient for a CasP6 polypeptide of the present disclosure and/or a nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure and/or a CasP6 guide RNA of the present disclosure is referred to as a "host cell" or a "target cell." A host cell or a target cell can be a recipient of a CasP6 system of the present disclosure. A host cell or a target cell can be a recipient of a CasP6 RNP of the present disclosure. A host cell or a target cell can be a recipient of a single component of a CasP6 system of the present disclosure.

[0203] Non-limiting examples of cells (target cells) include: a prokaryotic cell, eukaryotic cell, a bacterial cell, an archaeal cell, a cell of a single -cell eukaryotic organism, a protozoa cell, a cell from a plant (e.g., cells from plant crops, fruits, vegetables, grains, soy bean, corn, maize, wheat, seeds, tomatos, rice, cassava, sugarcane, pumpkin, hay, potatos, cotton, cannabis, tobacco,

flowering plants, conifers, gymnosperms, angiosperms, ferns, clubmosses, hornworts, liverworts, mosses, dicotyledons, monocotyledons, etc.), an algal cell, (e.g., Botryococcus braunii, Chlamydomonas reinhardtii, Nannochloropsis gaditana, Chlorella pyrenoidosa, Sargassum patens, C. agardh, and the like), seaweeds (e.g. kelp) a fungal cell (e.g., a yeast cell, a cell from a mushroom), an animal cell, a cell from an invertebrate animal (e.g., fruit fly, cnidarian, echinoderm, nematode, etc.), a cell from a vertebrate animal (e.g., fish, amphibian, reptile, bird, mammal), a cell from a mammal (e.g., an ungulate (e.g., a pig, a cow, a goat, a sheep); a rodent (e.g., a rat, a mouse); a non-human primate; a human; a feline (e.g., a cat); a canine (e.g., a dog); etc.), and the like. In some cases, the cell is a cell that does not originate from a natural organism (e.g., the cell can be a synthetically made cell; also referred to as an artificial cell).

[0204] A cell can be an in vitro cell (e.g., established cultured cell line). A cell can be an ex vivo cell (cultured cell from an individual). A cell can be and in vivo cell (e.g., a cell in an individual). A cell can be an isolated cell. A cell can be a cell inside of an organism. A cell can be an organism. A cell can be a cell in a cell culture (e.g., in vitro cell culture). A cell can be one of a collection of cells. A cell can be a prokaryotic cell or derived from a prokaryotic cell. A cell can be a bacterial cell or can be derived from a bacterial cell. A cell can be an archaeal cell or derived from an archaeal cell. A cell can be a eukaryotic cell or derived from a eukaryotic cell. A cell can be a plant cell or derived from a plant cell. A cell can be an animal cell or derived from an animal cell. A cell can be an invertebrate cell or derived from an invertebrate cell. A cell can be a vertebrate cell or derived from a vertebrate cell. A cell can be a mammalian cell or derived from a mammalian cell. A cell can be a rodent cell or derived from a rodent cell. A cell can be a human cell or derived from a human cell. A cell can be a microbe cell or derived from a microbe cell. A cell can be a fungi cell or derived from a fungi cell. A cell can be an insect cell. A cell can be an arthropod cell. A cell can be a protozoan cell. A cell can be a helminth cell.

[0205] Suitable cells include a stem cell (e.g. an embryonic stem (ES) cell, an induced pluripotent stem (iPS) cell; a germ cell (e.g., an oocyte, a sperm, an oogonia, a spermatogonia, etc.); a somatic cell, e.g. a fibroblast, an oligodendrocyte, a glial cell, a hematopoietic cell, a neuron, a muscle cell, a bone cell, a hepatocyte, a pancreatic cell, etc.

[0206] Suitable cells include human embryonic stem cells, fetal cardiomyocytes, myofibroblasts, mesenchymal stem cells, autotransplated expanded cardiomyocytes, adipocytes, totipotent cells, pluripotent cells, blood stem cells, myoblasts, adult stem cells, bone marrow cells, mesenchymal cells, embryonic stem cells, parenchymal cells, epithelial cells, endothelial cells, mesothelial cells, fibroblasts, osteoblasts, chondrocytes, exogenous cells, endogenous cells, stem cells, hematopoietic stem cells, bone- marrow derived progenitor cells, myocardial cells, skeletal cells,

fetal cells, undifferentiated cells, multi- potent progenitor cells, unipotent progenitor cells, monocytes, cardiac myoblasts, skeletal myoblasts, macrophages, capillary endothelial cells, xenogenic cells, allogenic cells, and post-natal stem cells.

[0207] In some cases, the cell is an immune cell, a neuron, an epithelial cell, and endothelial cell, or a stem cell. In some cases, the immune cell is a T cell, a B cell, a monocyte, a natural killer cell, a dendritic cell, or a macrophage. In some cases, the immune cell is a cytotoxic T cell. In some cases, the immune cell is a helper T cell. In some cases, the immune cell is a regulatory T cell (Treg).

[0208] In some cases, the cell is a stem cell. Stem cells include adult stem cells. Adult stem cells are also referred to as somatic stem cells.

[0209] Adult stem cells are resident in differentiated tissue, but retain the properties of self-renewal and ability to give rise to multiple cell types, usually cell types typical of the tissue in which the stem cells are found. Numerous examples of somatic stem cells are known to those of skill in the art, including muscle stem cells; hematopoietic stem cells; epithelial stem cells; neural stem cells; mesenchymal stem cells; mammary stem cells; intestinal stem cells; mesodermal stem cells; endothelial stem cells; olfactory stem cells; neural crest stem cells; and the like.

[0210] Stem cells of interest include mammalian stem cells, where the term "mammalian" refers to any animal classified as a mammal, including humans; non-human primates; domestic and farm animals; and zoo, laboratory, sports, or pet animals, such as dogs, horses, cats, cows, mice, rats, rabbits, etc. In some cases, the stem cell is a human stem cell. In some cases, the stem cell is a rodent (e.g., a mouse; a rat) stem cell. In some cases, the stem cell is a non-human primate stem cell. Stem cells can express one or more stem cell markers, e.g., SOX9, KRT19, KRT7, LGR5, CA9, FXYD2, CDH6, CLDN18, TSPAN8, BPIFB1, OLFM4, CDH17, and PPARGC1A.

[0211] In some embodiments, the stem cell is a hematopoietic stem cell (HSC). HSCs are mesoderm-derived cells that can be isolated from bone marrow, blood, cord blood, fetal liver and yolk sac. HSCs are characterized as $CD34^+$ and CD3 . HSCs can repopulate the erythroid, neutrophil- macrophage, megakaryocyte and lymphoid hematopoietic cell lineages in vivo. In vitro, HSCs can be induced to undergo at least some self-renewing cell divisions and can be induced to differentiate to the same lineages as is seen in vivo. As such, HSCs can be induced to differentiate into one or more of erythroid cells, megakaryocytes, neutrophils, macrophages, and lymphoid cells.

[0212] In other embodiments, the stem cell is a neural stem cell (NSC). NSCs are capable of differentiating into neurons, and glia (including oligodendrocytes, and astrocytes). A neural stem cell is a multipotent stem cell which is capable of multiple divisions, and under specific

conditions can produce daughter cells which are neural stem cells, or neural progenitor cells that can be neuroblasts or glioblasts, e.g., cells committed to become one or more types of neurons and glial cells respectively. Methods of obtaining NSCs are known in the art.

[0213] In other embodiments, the stem cell is a mesenchymal stem cell (MSC). MSCs originally derived from the embryonal mesoderm and isolated from adult bone marrow, can differentiate to form muscle, bone, cartilage, fat, marrow stroma, and tendon. Methods of isolating MSC are known in the art; and any known method can be used to obtain MSC. See, e.g., U.S. Pat. No. 5,736,396, which describes isolation of human MSC.

[0214] A cell is in some cases a plant cell. A plant cell can be a cell of a monocotyledon. A cell can be a cell of a dicotyledon. For example, the cell can be a cell of a major agricultural plant, e.g., Barley, Beans (Dry Edible), Canola, Corn, Cotton (Pima), Cotton (Upland), Flaxseed, Hay (Alfalfa), Hay (Non- Alfalfa), Oats, Peanuts, Rice, Sorghum, Soybeans, Sugarbeets, Sugarcane, Sunflowers (Oil), Sunflowers (Non-Oil), Sweet Potatoes, Tobacco (Burley), Tobacco (Flue-cured), Tomatoes, Wheat (Durum), Wheat (Spring), Wheat (Winter), and the like. As another example, the cell is a cell of a vegetable crops which include but are not limited to, e.g., alfalfa sprouts, aloe leaves, arrow root, arrowhead, artichokes, asparagus, bamboo shoots, banana flowers, bean sprouts, beans, beet tops, beets, bittermelon, bok choy, broccoli, broccoli rabe (rappini), brussels sprouts, cabbage, cabbage sprouts, cactus leaf (nopales), calabaza, cardoon, carrots, cauliflower, celery, chayote, Chinese artichoke (crosnes), Chinese cabbage, Chinese celery, Chinese chives, choy sum, chrysanthemum leaves (tung ho), collard greens, corn stalks, corn-sweet, cucumbers, daikon, dandelion greens, dasheen, dau mue (pea tips), donqua (winter melon), eggplant, endive, escarole, fiddle head ferns, field cress, frisee, gai choy (chinese mustard), gailon, galanga (siam, thai ginger), garlic, ginger root, gobo, greens, hanover salad greens, huauzontle, jerusalem artichokes, jicama, kale greens, kohlrabi, lamb's quarters (quilete), lettuce (bibb), lettuce (boston), lettuce (boston red), lettuce (green leaf), lettuce (iceberg), lettuce (lolla rossa), lettuce (oak leaf - green), lettuce (oak leaf - red), lettuce (processed), lettuce (red leaf), lettuce (romaine), lettuce (ruby romaine), lettuce (russian red mustard), linkok, lo bok, long beans, lotus root, mache, maguey (agave) leaves, malanga, mesculin mix, mizuna, moap (smooth luffa), moo, moqua (fuzzy squash), mushrooms, mustard, nagaimo, okra, ong choy, onions green, opo (long squash), ornamental corn, ornamental gourds, parsley, parsnips, peas, peppers (bell type), peppers, pumpkins, radicchio, radish sprouts, radishes, rape greens, rape greens, rhubarb, romaine (baby red), rutabagas, salicornia (sea bean), sinqua (angled/ridged luffa), spinach, squash, straw bales, sugarcane, sweet potatoes, swiss chard, tamarindo, taro, taro leaf, taro shoots, tatsoi, tepeguaje (guaje), tindora, tomatillos, tomatoes, tomatoes (cherry), tomatoes (grape

type), tomatoes (plum type), tumeric, turnip tops greens, turnips, water chestnuts, yampi, yams (names), yu choy, yuca (cassava), and the like.

[0215] A cell is in some cases an arthropod cell. For example, the cell can be a cell of a suborder, a family, a sub-family, a group, a sub-group, or a species of, e.g., Chelicerata, Myriapodia, Hexipodia, Arachnida, Insecta, Archaeognatha, Thysanura, Palaeoptera, Ephemeroptera, Odonata, Anisoptera, Zygoptera, Neoptera, Exopterygota, Plecoptera, Embioptera, Orthoptera, Zoraptera, Dermaptera, Dictyoptera, Notoptera, Grylloblattidae, Mantophasmatidae, Phasmatodea, Blattaria, Isoptera, Mantodea, Parapneuroptera, Psocoptera, Thysanoptera, Phthiraptera, Hemiptera, Endopterygota or Holometabola, Hymenoptera, Coleoptera, Strepsiptera, Raphidioptera, Megaloptera, Neuroptera, Mecoptera, Siphonaptera, Diptera, Trichoptera, or Lepidoptera.

[0216] A cell is in some cases an insect cell. For example, in some cases, the cell is a cell of a mosquito, a grasshopper, a true bug, a fly, a flea, a bee, a wasp, an ant, a louse, a moth, or a beetle.

[0217] The present disclosure provides a kit comprising a CasP6 system of the present disclosure, or a component of a CasP6 system of the present disclosure. A kit of the present disclosure can comprise: a) a CasP6 polypeptide of the present disclosure and a CasP6 guide RNA; b) a CasP6 polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; c) a CasP6 fusion polypeptide of the present disclosure and a CasP6 guide RNA; d) a CasP6 fusion polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; e) an mRNA encoding a CasP6 polypeptide of the present disclosure; and a CasP6 guide RNA; f) an mRNA encoding a CasP6 polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; g) an mRNA encoding a CasP6 fusion polypeptide of the present disclosure; and a CasP6 guide RNA; h) an mRNA encoding a CasP6 fusion polypeptide of the present disclosure, a CasP6 guide RNA, and a donor template nucleic acid; i) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure and a nucleotide sequence encoding a CasP6 guide RNA; j) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, a nucleotide sequence encoding a CasP6 guide RNA, and a nucleotide sequence encoding a donor template nucleic acid; k) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure and a nucleotide sequence encoding a CasP6 guide RNA; 1) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, a nucleotide sequence encoding a CasP6 guide RNA, and a nucleotide sequence encoding a donor

template nucleic acid; m) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; n) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; and a donor template nucleic acid; o) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; p) a first recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, and a second recombinant expression vector comprising a nucleotide sequence encoding a CasP6 guide RNA; and a donor template nucleic acid; q) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure, a nucleotide sequence encoding a first CasP6 guide RNA, and a nucleotide sequence encoding a second CasP6 guide RNA; or r) a recombinant expression vector comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure, a nucleotide sequence encoding a first CasP6 guide RNA, and a nucleotide sequence encoding a second CasP6 guide RNA; or some variation of one of (a) through (r).

[0218] A kit of the present disclosure can comprise: a) a component, as described above, of a CasP6 system of the present disclosure, or can comprise a CasP6 system of the present disclosure; and b) one or more additional reagents, e.g., i) a buffer; ii) a protease inhibitor; iii) a nuclease inhibitor; iv) a reagent required to develop or visualize a detectable label; v) a positive and/or negative control target DNA or RNA; vi) a positive and/or negative control CasP6 guide RNA; and the like. A kit of the present disclosure can comprise: a) a component, as described above, of a CasP6 system of the present disclosure, or can comprise a CasP6 system of the present disclosure; and b) a therapeutic agent. A kit of the present disclosure can comprise a recombinant expression vector comprising: a) an insertion site for inserting a nucleic acid comprising a nucleotide sequence encoding a portion of a CasP6 guide RNA that hybridizes to a target nucleotide sequence in a target nucleic acid; and b) a nucleotide sequence encoding the CasP6 -binding portion of a CasP6 guide RNA. A kit of the present disclosure can comprise a recombinant expression vector comprising: a) an insertion site for inserting a nucleic acid comprising a nucleotide sequence encoding a portion of a CasP6 guide RNA that hybridizes to a target nucleotide sequence in a target nucleic acid; b) a nucleotide sequence encoding the CasP6-

binding portion of a CasP6 guide RNA; and c) a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure.

[0219] A CasP6 polypeptide of the present disclosure, or a CasP6 fusion polypeptide of the present disclosure, finds use in a variety of methods (e.g., in combination with a CasP6 guide RNA and in some cases further in combination with a donor template). For example, a CasP6 polypeptide of the present disclosure can be used to (i) modify (e.g., cleave, e.g., nick; methylate; etc.) target nucleic acid (DNA or RNA; single stranded or double stranded); (ii) modulate transcription of a target nucleic acid; (iii) label a target nucleic acid; (iv) bind a target nucleic acid (e.g., for purposes of isolation, labeling, imaging, tracking, etc.); (v) modify a polypeptide (e.g., a histone) associated with a target nucleic acid; and the like. Thus, the present disclosure provides a method of modifying a target nucleic acid. In some cases, a method of the present disclosure for modifying a target nucleic acid comprises contacting the target nucleic acid with: a) a CasP6 polypeptide of the present disclosure; and b) one or more (e.g., two) CasP6 guide RNAs. In some cases, a method of the present disclosure for modifying a target nucleic acid comprises contacting the target nucleic acid with: a) a CasP6 polypeptide of the present disclosure; b) a CasP6 guide RNA; and c) a donor nucleic acid (e.g, a donor template). In some cases, the contacting step is carried out in a cell in vitro. In some cases, the contacting step is carried out in a cell in vivo. In some cases, the contacting step is carried out in a cell ex vivo.

[0220] Because a method that uses a CasP6 polypeptide includes binding of the CasP6 polypeptide to a particular region in a target nucleic acid (by virtue of being targeted there by an associated CasP6 guide RNA), the methods are generally referred to herein as methods of binding (e.g., a method of binding a target nucleic acid). However, it is to be understood that in some cases, while a method of binding may result in nothing more than binding of the target nucleic acid, in other cases, the method can have different final results (e.g., the method can result in modification of the target nucleic acid, e.g., cleavage/me thylation/etc, modulation of transcription from the target nucleic acid; modulation of translation of the target nucleic acid; genome editing; modulation of a protein associated with the target nucleic acid; isolation of the target nucleic acid; etc.).

[0221] For examples of suitable methods, see, for example, Jinek et al., Science. 2012 Aug 17;337(6096):816-21 ; Chylinski et al., RNA Biol. 2013 May;10(5):726-37; Ma et al., Biomed Res Int. 2013;2013:270805; Hou et al., Proc Natl Acad Sci U S A. 2013 Sep 24;110(39):15644-9; Jinek et al., Elife. 2013;2:e00471 ; Pattanayak et al., Nat Biotechnol. 2013 Sep;31(9):839-43; Qi et al, Cell. 2013 Feb 28;152(5): 1173-83; Wang et al., Cell. 2013 May 9;153(4):910-8; Auer et al, Genome Res. 2013 Oct 31 ; Chen et al., Nucleic Acids Res. 2013 Nov l ;41(20):el9; Cheng et al.,

Cell Res. 2013 Oct;23(10): 1163-71 ; Cho et al., Genetics. 2013 Nov;195(3): 1177-80; DiCarlo et al., Nucleic Acids Res. 2013 Apr;41(7):4336- 43; Dickinson et al., Nat Methods. 2013 Oct;10(10): 1028-34; Ebina et al., Sci Rep. 2013;3:2510; Fujii et al, Nucleic Acids Res. 2013 Nov l ;41(20):el87; Hu et al., Cell Res. 2013 Nov;23(l l): 1322-5; Jiang et al., Nucleic Acids Res. 2013 Nov l ;41(20):el88; Larson et al, Nat Protoc. 2013 Nov;8(l l):2180-96; Mali et. at., Nat Methods. 2013 Oct;10(10):957-63; Nakayama et al., Genesis. 2013 Dec;51(12):835-43; Ran et al., Nat Protoc. 2013 Nov;8(l l):2281-308; Ran et al., Cell. 2013 Sep 12;154(6): 1380-9; Upadhyay et al., G3 (Bethesda). 2013 Dec 9;3(12):2233-8; Walsh et al., Proc Natl Acad Sci U S A. 2013 Sep 24;110(39): 15514-5; Xie et al., Mol Plant. 2013 Oct 9; Yang et al., Cell. 2013 Sep 12;154(6): 1370-9; L. B. Harrington et al., Science 10.1126/science.aav4294 (2018), and U.S. patents and patent applications: 8,906,616; 8,895,308; 8,889,418; 8,889,356; 8,871,445; 8,865,406; 8,795,965; 8,771,945; 8,697,359; 20140068797; 20140170753; 20140179006; 20140179770; 20140186843; 20140186919; 20140186958; 20140189896; 20140227787; 20140234972; 20140242664; 20140242699; 20140242700; 20140242702; 20140248702; 20140256046; 20140273037; 20140273226; 20140273230; 20140273231 ; 20140273232; 20140273233; 20140273234; 20140273235; 20140287938; 20140295556; 20140295557; 20140298547; 20140304853; 20140309487; 20140310828; 20140310830; 20140315985; 20140335063; 20140335620; 20140342456; 20140342457; 20140342458; 20140349400; 20140349405; 20140356867; 20140356956; 20140356958; 20140356959; 20140357523; 20140357530; 20140364333; and 20140377868; each of which is hereby incorporated by reference in its entirety.

[0222] For example, the present disclosure provides (but is not limited to) methods of cleaving a target nucleic acid; methods of editing a target nucleic acid; methods of modulating transcription from a target nucleic acid; methods of isolating a target nucleic acid, methods of binding a target nucleic acid, methods of imaging a target nucleic acid, methods of modifying a target nucleic acid, and the like.

[0223] As used herein, the terms/phrases "contact a target nucleic acid" and "contacting a target nucleic acid", for example, with a CasP6 polypeptide or with a CasP6 fusion polypeptide, etc., encompass all methods for contacting the target nucleic acid. For example, a CasP6 polypeptide can be provided to a cell as protein, RNA (encoding the CasP6 polypeptide), or DNA (encoding the CasP6 polypeptide); while a CasP6 guide RNA can be provided as a guide RNA or as a nucleic acid encoding the guide RNA. As such, when, for example, performing a method in a cell (e.g., inside of a cell in vitro, inside of a cell in vivo, inside of a cell ex vivo), a method that includes contacting the target nucleic acid encompasses the introduction into the cell of any or all

of the components in their active/final state (e.g., in the form of a protein(s) for CasP6 polypeptide; in the form of a protein for a CasP6 fusion polypeptide; in the form of an RNA in some cases for the guide RNA), and also encompasses the introduction into the cell of one or more nucleic acids encoding one or more of the components (e.g., nucleic acid(s) comprising nucleotide sequence(s) encoding a CasP6 polypeptide or a CasP6 fusion polypeptide, nucleic acid(s) comprising nucleotide sequence(s) encoding guide RNA(s), nucleic acid comprising a nucleotide sequence encoding a donor template, and the like). Because the methods can also be performed in vitro outside of a cell, a method that includes contacting a target nucleic acid, (unless otherwise specified) encompasses contacting outside of a cell in vitro, inside of a cell in vitro, inside of a cell in vivo, inside of a cell ex vivo, etc.

[0224] In some cases, a method of the present disclosure for modifying a target nucleic acid comprises contacting a target nucleic acid with a CasP6 polypeptide of the present disclosure, or with a CasP6 fusion polypeptide of the present disclosure. In some cases, abmethod of the present disclosure for modifying a target nucleic acid comprises contacting a target nucleic acid with a CasP6 polypeptide and a CasP6 guide RNA. In some cases, abmethod of the present disclosure for modifying a target nucleic acid comprises contacting a target nucleic acid with a CasP6 polypeptidel, a first CasP6 guide RNA, and a second CasP6 guide RNA In some cases, a method of the present disclosure for modifying a target nucleic acid comprises contacting a target nucleic acid with a CasP6 polypeptide of the present disclosure and a CasP6 guide RNA and a donor DNA template.

[0225] A CasP6 polypeptide of the present disclosure, or a CasP6 fusion polypeptide of the present disclosure, when bound to a CasP6 guide RNA, can bind to a target nucleic acid, and in some cases, can bind to and modify a target nucleic acid. A target nucleic acid can be any nucleic acid (e.g., DNA, RNA), can be double stranded or single stranded, can be any type of nucleic acid (e.g., a chromosome (genomic DNA), derived from a chromosome, chromosomal DNA, plasmid, viral, extracellular, intracellular, mitochondrial, chloroplast, linear, circular, etc.) and can be from any organism (e.g., as long as the CasP6 guide RNA comprises a nucleotide sequence that hybridizes to a target sequence in a target nucleic acid, such that the target nucleic acid can be targeted).

[0226] A target nucleic acid can be DNA or RNA. A target nucleic acid can be double stranded (e.g., dsDNA, dsRNA) or single stranded (e.g., ssRNA, ssDNA). In some cases, a target nucleic acid is single stranded. In some cases, a target nucleic acid is a single stranded RNA (ssRNA). In some cases, a target ssRNA (e.g., a target cell ssRNA, a viral ssRNA, etc.) is selected from: mRNA, rRNA, tRNA, non- coding RNA (ncRNA), long non-coding RNA (IncRNA), and

microRNA (miRNA). In some cases, a target nucleic acid is a single stranded DNA (ssDNA) (e.g., a viral DNA). As noted above, in some cases, a target nucleic acid is single stranded.

[0227] A target nucleic acid can be located anywhere, for example, outside of a cell in vitro, inside of a cell in vitro, inside of a cell in vivo, inside of a cell ex vivo. Suitable target cells (which can comprise target nucleic acids such as genomic DNA) include, but are not limited to: a bacterial cell; an archaeal cell; a cell of a single-cell eukaryotic organism; a plant cell; an algal cell, e.g., Botryococcus braunii, Chlamydomonas reinhardtii, Nannochloropsis gaditana, Chlorella pyrenoidosa, Sargassum patens, C. agardh, and the like; a fungal cell (e.g., a yeast cell); an animal cell; a cell from an invertebrate animal (e.g. fruit fly, a cnidarian, an echinoderm, a nematode, etc.); a cell of an insect (e.g., a mosquito; a bee; an agricultural pest; etc.); a cell of an arachnid (e.g., a spider; a tick; etc.); a cell from a vertebrate animal (e.g., a fish, an amphibian, a reptile, a bird, a mammal); a cell from a mammal (e.g., a cell from a rodent; a cell from a human; a cell of a non-human mammal; a cell of a rodent (e.g., a mouse, a rat); a cell of a lagomorph (e.g., a rabbit); a cell of an ungulate (e.g., a cow, a horse, a camel, a llama, a vicuna, a sheep, a goat, etc.); a cell of a marine mammal (e.g., a whale, a seal, an elephant seal, a dolphin, a sea lion; etc.) and the like. Any type of cell may be of interest (e.g. a stem cell, e.g. an embryonic stem (ES) cell, an induced pluripotent stem (iPS) cell, a germ cell (e.g., an oocyte, a sperm, an oogonia, a spermatogonia, etc.), an adult stem cell, a somatic cell, e.g. a fibroblast, a hematopoietic cell, a neuron, a muscle cell, a bone cell, a hepatocyte, a pancreatic cell; an in vitro or in vivo embryonic cell of an embryo at any stage, e.g., a 1-cell, 2-cell, 4-cell, 8-cell, etc. stage zebrafish embryo; etc.).

[0228] Cells may be from established cell lines or they may be primary cells, where "primary cells", "primary cell lines", and "primary cultures" are used interchangeably herein to refer to cells and cells cultures that have been derived from a subject and allowed to grow in vitro for a limited number of passages, i.e. splittings, of the culture. For example, primary cultures are cultures that may have been passaged 0 times, 1 time, 2 times, 4 times, 5 times, 10 times, or 15 times, but not enough times go through the crisis stage. Typically, the primary cell lines are maintained for fewer than 10 passages in vitro. Target cells can be unicellular organisms and/or can be grown in culture. If the cells are primary cells, they may be harvest from an individual by any convenient method. For example, leukocytes may be conveniently harvested by apheresis, leukocytapheresis, density gradient separation, etc., while cells from tissues such as skin, muscle, bone marrow, spleen, liver, pancreas, lung, intestine, stomach, etc. can be conveniently harvested by biopsy.

[0229] In some of the above applications, the subject methods may be employed to induce target nucleic acid cleavage, target nucleic acid modification, and/or to bind target nucleic acids (e.g., for visualization, for collecting and/or analyzing, etc.) in mitotic or post-mitotic cells in vivo and/or ex vivo and/or in vitro (e.g., to disrupt production of a protein encoded by a targeted mRNA, to cleave or otherwise modify target DNA, to genetically modify a target cell, and the like). Because the guide RNA provides specificity by hybridizing to target nucleic acid, a mitotic and/or post-mitotic cell of interest in the disclosed methods may include a cell from any organism (e.g. a bacterial cell, an archaeal cell, a cell of a single-cell eukaryotic organism, a plant cell, an algal cell, e.g., Botryococcus braunii, Chlamydomonas reinhardtii, Nannochloropsis gaditana, Chlorella pyrenoidosa, Sargassum patens, C. agardh, and the like, a fungal cell (e.g., a yeast cell), an animal cell, a cell from an invertebrate animal (e.g. fruit fly, cnidarian, echinoderm, nematode, etc.), a cell from a vertebrate animal (e.g., fish, amphibian, reptile, bird, mammal), a cell from a mammal, a cell from a rodent, a cell from a human, etc.). In some cases, a subject CasP6 protein (and/or nucleic acid encoding the protein such as DNA and/or RNA), and/or CasP6 guide RNA (and/or a DNA encoding the guide RNA), and/or donor template, and/or RNP can be introduced into an individual (i.e., the target cell can be in vivo) (e.g., a mammal, a rat, a mouse, a pig, a primate, a non-human primate, a human, etc.). In some case, such an administration can be for the purpose of treating and/or preventing a disease, e.g., by editing the genome of targeted cells.

[0230] Plant cells include cells of a monocotyledon, and cells of a dicotyledon. The cells can be root cells, leaf cells, cells of the xylem, cells of the phloem, cells of the cambium, apical meristem cells, parenchyma cells, collenchyma cells, sclerenchyma cells, and the like. Plant cells include cells of agricultural crops such as wheat, corn, rice, sorghum, millet, soybean, etc. Plant cells include cells of agricultural fruit and nut plants, e.g., plant that produce apricots, oranges, lemons, apples, plums, pears, almonds, etc.

[0231] Non-limiting examples of cells can be found in the section “Modified host cells”.

[0232] Guided by a CasP6 single guide RNA, a CasP6 protein in some cases generates site-specific double strand breaks (DSBs) or single strand breaks (SSBs) (e.g., when the CasP6 protein is a nickase variant) within double-stranded DNA (dsDNA) target nucleic acids, which are repaired either by non-homologous end joining (NHEJ) or homology-directed recombination (HDR).

[0233] In some cases, contacting a target DNA (with a CasP6 protein and a CasP6 guide RNA) occurs under conditions that are permissive for nonhomologous end joining or homology-directed repair. Thus, in some cases, a subject method includes contacting the target DNA with a donor polynucleotide (e.g., by introducing the donor polynucleotide into a cell), wherein the donor

polynucleotide, a portion of the donor polynucleotide, a copy of the donor polynucleotide, or a portion of a copy of the donor polynucleotide integrates into the target DNA. In some cases, the method does not comprise contacting a cell with a donor polynucleotide, and the target DNA is modified such that nucleotides within the target DNA are deleted.

[0234] In some cases, CasP6 guide RNA (or DNA encoding same) and a CasP6 protein (or a nucleic acid encoding same, such as an RNA or a DNA, e.g, one or more expression vectors) are coadministered (e.g., contacted with a target nucleic acid, administered to cells, etc.) with a donor polynucleotide sequence that includes at least a segment with homology to the target DNA sequence, the subject methods may be used to add, i.e. insert or replace, nucleic acid material to a target DNA sequence (e.g. to "knock in" a nucleic acid, e.g., one that encodes for a protein, an siRNA, an miRNA, etc.), to add a tag (e.g., 6xHis, a fluorescent protein (e.g., a green fluorescent protein; a yellow fluorescent protein, etc.), hemagglutinin (HA), FLAG, etc.), to add a regulatory sequence to a gene (e.g. promoter, polyadenylation signal, internal ribosome entry sequence (IRES), 2A peptide, start codon, stop codon, splice signal, localization signal, etc.), to modify a nucleic acid sequence (e.g., introduce a mutation, remove a disease causnig mutation by introducing a correct sequence), and the like. As such, a complex comprising a CasP6 guide RNA and CasP6 protein is useful in any in vitro or in vivo application in which it is desirable to modify DNA in a site-specific, i.e. "targeted", way, for example gene knock-out, gene knock-in, gene editing, gene tagging, etc., as used in, for example, gene therapy, e.g. to treat a disease or as an antiviral, antipathogenic, or anticancer therapeutic, the production of genetically modified organisms in agriculture, the large scale production of proteins by cells forntherapeutic, diagnostic, or research purposes, the induction of iPS cells, biological research, the targeting of genes of pathogens for deletion or replacement, etc.

[0235] In applications in which it is desirable to insert a polynucleotide sequence into the genome where a target sequence is cleaved, a donor polynucleotide (a nucleic acid comprising a donor sequence) can also be provided to the cell. By a "donor sequence" or "donor polynucleotide" or “DNA donor template” or "donor template" it is meant a nucleic acid sequence to be inserted at the site cleaved by the CasP6 protein (e.g., after dsDNA cleavage, after nicking a target DNA, after dual nicking a target DNA, and the like). The donor polynucleotide or DNA donor template can contain sufficient homology to a genomic sequence at the target site, e.g. 70%, 80%, 85%, 90%, 95%, or 100% homology with the nucleotide sequences flanking the target site, e.g. within about 50 bases or less of the target site, e.g. within about 30 bases, within about 15 bases, within about 10 bases, within about 5 bases, or immediately flanking the target site, to support homology-directed repair between it and the genomic sequence to which it bears

homology. Approximately 25, 50, 100, or 200 nucleotides, or more than 200 nucleotides, of sequence homology between a donor or DNA donor template and a genomic sequence (or any integral value between 10 and 200 nucleotides, or more) can support homology-directed repair. Donor polynucleotides or DNA donor template can be of any length, e.g. 10 nucleotides or more, 50 nucleotides or more, 100 nucleotides or more, 250 nucleotides or more, 500 nucleotides or more, 1000 nucleotides or more, 5000 nucleotides or more, etc.

[0236] The donor sequence or "DNA donor template"is typically not identical to the genomic sequence that it replaces. Rather, the donor sequence or DNA donor template may contain at least one or more single base changes, insertions, deletions, inversions or rearrangements with respect to the genomic sequence, so long as sufficient homology is present to support homology-directed repair (e.g., for gene correction, e.g., to convert a disease-causing base pair or a non disease-causing base pair). In some embodiments, the donor sequence or DNA donor template comprises a nonhomologous sequence flanked by two regions of homology, such that homology-directed repair between the target DNA region and the two flanking sequences results in insertion of the non-homologous sequence at the target region. Donor sequences or DNA donor templates may also comprise a vector backbone containing sequences that are not homologous to the DNA region of interest and that are not intended for insertion into the DNA region of interest. Generally, the homologous region(s) of a donor sequence or DNA donor template will have at least 50% sequence identity to a genomic sequence with which recombination is desired. In certain embodiments, 60%, 70%, 80%, 90%, 95%, 98%, 99%, or 99.9% sequence identity is present. Any value between 1% and 100% sequence identity can be present, depending upon the length of the donor polynucleotide or DNA donor template.

[0237] The donor sequence or DNA donor template may comprise certain sequence differences as compared to the genomic sequence, e.g. restriction sites, nucleotide polymorphisms, selectable markers (e.g., drug resistance genes, fluorescent proteins, enzymes etc.), etc., which may be used to assess for successful insertion of the donor sequence or DNA donor template at the cleavage site or in some cases may be used for other purposes (e.g., to signify expression at the targeted genomic locus). In some cases, if located in a coding region, such nucleotide sequence differences will not change the amino acid sequence, or will make silent amino acid changes (i.e., changes which do not affect the structure or function of the protein). Alternatively, these sequences differences may include flanking recombination sequences such as FLPs, loxP sequences, or the like, that can be activated at a later time for removal of the marker sequence.

[0238] In some cases, the donor sequence or DNA donor template is provided to the cell as single-stranded DNA. In some cases, the donor sequence or DNA donor template is provided to

the cell as double-stranded DNA. It may be introduced into a cell in linear or circular form. If introduced in linear form, the ends of the donor sequence or DNA donor template may be protected (e.g., from exonucleolytic degradation) by any convenient method and such methods are known to those of skill in the art. For example, one or more dideoxynucleotide residues can be added to the 3' terminus of a linear molecule and/or self -complementary oligonucleotides can be ligated to one or both ends. See, for example, Chang et al. (1987) Proc. Natl. Acad Sci USA 84:4959-4963; Nehls et al. (1996) Science 272:886-889. Additional methods for protecting exogenous polynucleotides from degradation include, but are not limited to, addition of terminal amino group(s) and the use of modified internucleotide linkages such as, for example, phosphorothioates, phosphor amidates, and O-methyl ribose or deoxyribose residues. As an alternative to protecting the termini of a linear donor sequence or DNA donor template, additional lengths of sequence may be included outside of the regions of homology that can be degraded without impacting recombination. A donor sequence or DNA donor template can be introduced into a cell as part of a vector molecule having additional sequences such as, for example, replication origins, promoters and genes encoding antibiotic resistance. Moreover, donor sequences or DNA donor template can be introduced as naked nucleic acid, as nucleic acid complexed with an agent such as a liposome or poloxamer, or can be delivered by viruses (e.g., adenovirus, AAV), as described elsewhere herein for nucleic acids encoding a CasP6 guide RNA and/or a CasP6 fusion polypeptide and/or donor polynucleotide.

[0239] As described above, in some cases, a nucleic acid (e.g., a recombinant expression vector) of the present disclosure (e.g., a nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure; a nucleic acid comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure; etc.), is used as a transgene to generate a transgenic non- human organism that produces a CasP6 polypeptide, or a CasP6 fusion polypeptide, of the present disclosure. The present disclosure provides a transgenic non-human organism comprising a nucleotide sequence encoding a CasP6 polypeptide, or a CasP6 fusion polypeptide, of the present disclosure.

[0240] The present disclosure provides a transgenic non-human animal, which animal comprises a transgene comprising a nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide or a CasP6 fusion polypeptide. In some embodiments, the genome of the transgenic non-human animal comprises a nucleotide sequence encoding a CasP6 polypeptide or a CasP6 fusion polypeptide, of the present disclosure. In some cases, the transgenic non-human animal is homozygous for the genetic modification. In some cases, the transgenic non-human animal is heterozygous for the genetic modification. In some embodiments, the transgenic non-human

animal is a vertebrate, for example, a fish (e.g., salmon, trout, zebra fish, gold fish, puffer fish, cave fish, etc.), an amphibian (frog, newt, salamander, etc.), a bird (e.g., chicken, turkey, etc.), a reptile (e.g., snake, lizard, etc.), a non-human mammal (e.g., an ungulate, e.g., a pig, a cow, a goat, a sheep, etc.; a lagomorph (e.g., a rabbit); a rodent (e.g., a rat, a mouse); a non-human primate; etc.), etc. In some cases, the transgenic non-human animal is an invertebrate. In some cases, the transgenic non-human animal is an insect (e.g., a mosquito; an agricultural pest; etc.). In some cases, the transgenic non-human animal is an arachnid.

[0241] Nucleotide sequences encoding a CasP6 polypeptide or a CasP6 fusion polypeptide, of the present disclosure can be under the control of (i.e., operably linked to) an unknown promoter (e.g., when the nucleic acid randomly integrates into a host cell genome) or can be under the control of (i.e., operably linked to) a known promoter. Suitable known promoters can be any known promoter and include constitutively active promoters (e.g., CMV promoter), inducible promoters (e.g., heat shock promoter, tetracycline-regulated promoter, steroid-regulated promoter, metal-regulated promoter, estrogen receptor- regulated promoter, etc.), spatially restricted and/or temporally restricted promoters (e.g., a tissue specific promoter, a cell type specific promoter, etc.).

[0242] As described above, in some cases, a nucleic acid (e.g., a recombinant expression vector) of the present disclosure (e.g., a nucleic acid comprising a nucleotide sequence encoding a CasP6 polypeptide of the present disclosure; a nucleic acid comprising a nucleotide sequence encoding a CasP6 fusion polypeptide of the present disclosure; etc.), is used as a transgene to generate a transgenic plant that produces a CasP6 polypeptide, or a CasP6 fusion polypeptide, of the present disclosure. The present disclosure provides a transgenic plant comprising a nucleotide sequence encoding a CasP6 polypeptide, or a CasP6 fusion polypeptide, of the present disclosure. In some embodiments, the genome of the transgenic plant comprises a subject nucleic acid. In some embodiments, the transgenic plant is homozygous for the genetic modification. In some embodiments, the transgenic plant is heterozygous for the genetic modification.

[0243] Methods of introducing exogenous nucleic acids into plant cells are well known in the art. Such plant cells are considered "transformed," as defined above. Suitable methods include viral infection (such as double stranded DNA viruses), transfection, conjugation, protoplast fusion, electroporation, particle gun technology, calcium phosphate precipitation, direct microinjection, silicon carbide whiskers technology, Agrobacterium-mediated transformation and the like. The choice of method is generally dependent on the type of cell being transformed and the circumstances under which the transformation is taking place (i.e. in vitro, ex vivo, or in vivo).

[0244] Transformation methods based upon the soil bacterium Agrobacterium tumefaciens are particularly useful for introducing an exogenous nucleic acid molecule into a vascular plant. The wild type form of Agrobacterium contains a Ti (tumor-inducing) plasmid that directs production of tumorigenic crown gall growth on host plants. Transfer of the tumor-inducing T-DNA region of the Ti plasmid to a plant genome requires the Ti plasmid-encoded virulence genes as well as T-DNA borders, which are a set of direct DNA repeats that delineate the region to be transferred. An Agrobacterium-based vector is a modified form of a Ti plasmid, in which the tumor inducing functions are replaced by the nucleic acid sequence of interest to be introduced into the plant host.

[0245] Agrobacterium-mediated transformation generally employs cointegrate vectors or binary vector systems, in which the components of the Ti plasmid are divided between a helper vector, which resides permanently in the Agrobacterium host and carries the virulence genes, and a shuttle vector, which contains the gene of interest bounded by T-DNA sequences. A variety of binary vectors is well known in the art and are commercially available, for example, from Clontech (Palo Alto, Calif.). Methods of coculturing Agrobacterium with cultured plant cells or wounded tissue such as leaf tissue, root explants, hypocotyledons, stem pieces or tubers, for example, also are well known in the art. See, e.g., Glick and Thompson, (eds.), Methods in Plant Molecular Biology and Biotechnology, Boca Raton, Fla.: CRC Press (1993).

[0246] Microprojectile-mediated transformation also can be used to produce a subject transgenic plant. This method, first described by Klein et al. (Nature 327:70-73 (1987)), relies on microprojectiles such as gold or tungsten that are coated with the desired nucleic acid molecule by precipitation with calcium chloride, spermidine or polyethylene glycol. The microprojectile particles are accelerated at high speed into an angiosperm tissue using a device such as the BIOLISTIC PD-1000 (Biorad; Hercules Calif.).

[0247] A nucleic acid of the present disclosure (e.g., a nucleic acid (e.g., a recombinant expression vector) comprising a nucleotide sequence encoding a CasP6 polypeptide, or a CasP6 fusion polypeptide, of the present disclosure) may be introduced into a plant in a manner such that the nucleic acid is able to enter a plant cell(s), e.g., via an in vivo or ex vivo protocol. By "in vivo," it is meant in the nucleic acid is administered to a living body of a plant e.g. infiltration. By "ex vivo" it is meant that cells or explants are modified outside of the plant, and then such cells or organs are regenerated to a plant. A number of vectors suitable for stable transformation of plant cells or for the establishment of transgenic plants have been described, including those described in Weissbach and Weissbach, (1989) Methods for Plant Molecular Biology Academic Press, and Gelvin et al., (1990) Plant Molecular Biology Manual, Kluwer Academic Publishers.

Specific examples include those derived from a Ti plasmid of Agrobacterium tumefaciens, as well as those disclosed by Herrera-Estrella et al. (1983) Nature 303: 209, Bevan (1984) Nucl Acid Res. 12: 8711-8721, Klee (1985) Bio/Technolo 3: 637-642. Alternatively, non- Ti vectors can be used to transfer the DNA into plants and cells by using free DNA delivery techniques. By using these methods transgenic plants such as wheat, rice (Christou (1991) Bio/Technology 9:957-9 and 4462) and corn (Gordon-Kamm (1990) Plant Cell 2: 603-618) can be produced. An immature embryo can also be a good target tissue for monocots for direct DNA delivery techniques by using the particle gun (Weeks et al. (1993) Plant Physiol 102: 1077-1084; Vasil (1993) Bio/Technolo 10: 667-674; Wan and Lemeaux (1994) Plant Physiol 104: 37-48 and for Agrobacterium-mediated DNA transfer (Ishida et al. (1996) Nature Biotech 14: 745-750). Exemplary methods for introduction of DNA into chloroplasts are biolistic bombardment, polyethylene glycol transformation of protoplasts, and microinjection (Danieli et al Nat. Biotechnol 16:345-348, 1998; Staub et al Nat. Biotechnol 18: 333-338, 2000; O'Neill et al Plant J. 3:729-738, 1993; Knoblauch et al Nat. Biotechnol 17: 906-909; U.S. Pat. Nos. 5,451,513, 5,545,817, 5,545,818, and 5,576,198; in Intl. Application No. WO 95/16783; and in Boynton et al., Methods in Enzymology 217: 510-536 (1993), Svab et al., Proc. Natl. Acad. Sci. USA 90: 913-917 (1993), and McBride et al., Proc. Natl. Acad. Sci. USA 91 : 7301-7305 (1994)). Any vector suitable for the methods of biolistic bombardment, polyethylene glycol transformation of protoplasts and microinjection will be suitable as a targeting vector for chloroplast transformation. Any double stranded DNA vector may be used as a transformation vector, especially when the method of introduction does not utilize Agrobacterium.

[0248] Plants which can be genetically modified include grains, forage crops, fruits, vegetables, oil seed crops, palms, forestry, and vines. Specific examples of plants which can be modified follow: maize, banana, peanut, field peas, sunflower, tomato, canola, tobacco, wheat, barley, oats, potato, soybeans, cotton, carnations, sorghum, lupin and rice.

[0249] The present disclosure provides transformed plant cells, tissues, plants and products that contain the transformed plant cells. A feature of the subject transformed cells, and tissues and products that include the same is the presence of a subject nucleic acid integrated into the genome, and production by plant cells of a CasP6 polypeptide, or a CasP6 fusion polypeptide, of the present disclosure.

[0250] Recombinant plant cells of the present invention are useful as populations of recombinant cells, or as a tissue, seed, whole plant, stem, fruit, leaf, root, flower, stem, tuber, grain, animal feed, a field of plants, and the like.

[0251] Nucleotide sequences encoding a CasP6 polypeptide, or a CasP6 fusion polypeptide, of the present disclosure can be under the control of (i.e., operably linked to) an unknown promoter (e.g., when the nucleic acid randomly integrates into a host cell genome) or can be under the control of (i.e., operably linked to) a known promoter. Suitable known promoters can be any known promoter and include constitutively active promoters, inducible promoters, spatially restricted and/or temporally restricted promoters, etc.

[0252] In some embodiments the disclosed CasP6 RNA guided nucleases can be used in systems and methods for detecting one or more specific target RNA molecules in a sample. Examples of target RNA molecule detection schemes that were implemented with distinct RNA guided nucleases are described in US20180274017, which is hereby incorporated by reference in its entirety. In certain embodiments, the methods and reagents (*e.g.,* RNA masking constructs) described in US20180274017 and incorporated herein by reference can be adapted for use with the CasP6 polypeptides, CasP6 fusion polypeptides, and CasP6 guide RNA molecules disclosed herein.

[0253] Guide RNAs for a CasP6 polypeptide or fusion polypeptide are designed to recognize a target RNA molecule having target sequences in a sample suspected of having the target RNA of interest. If the sample contains the target RNA molecule, binding of the target RNA molecule by the CasP6 guide RNA/CasP6 polypeptide or fusion polypeptide complex will trigger CasP6 polypeptide or fusion polypeptide's collateral nuclease activity (*e.g.,* cleavage of a single stranded RNA (ssRNA) that does not contain the target RNA sequences or contains one or more mismatches to the target sequence). Consequently, the RNA-based masking construct will produce an output following cleavage by the CasP6 polypeptide or fusion polypeptide collateral nuclease activity that can be assayed. Presence or absence of the output, therefore, indicates presence or absence of a RNA having the target sequence in the sample.

[0254] As used herein, a "masking construct" refers to a molecule that can be cleaved or otherwise deactivated by an activated CasP6 system described herein. The term "masking construct" may also be referred to in the alternative as a "detection construct." In certain example embodiments, the masking construct is an RNA-based masking construct. The RNA-based masking construct comprises an RNA element that is cleavable by a CasP6 protein. Cleavage of the RNA element releases agents or produces conformational changes that allow a detectable signal to be produced. Example constructs demonstrating how the RNA element may be used to prevent or mask generation of detectable signal are described below and embodiments comprise variants of the same. Prior to cleavage, or when the masking construct is in an "active" state, the masking construct blocks the generation or detection of a positive detectable signal. It will be

understood that in certain example embodiments a minimal background signal may be produced in the presence of an active RNA masking construct. A positive detectable signal may be any signal that can be detected using optical, fluorescent, chemiluminescent, electrochemical or other detection methods known in the art. The term "positive detectable signal" is used to differentiate from other detectable signals that may be detectable in the presence of the masking construct. For example, in certain embodiments a first signal may be detected when the masking agent is present (i.e. a negative detectable signal), which then converts to a second signal (e.g. the positive detectable signal) upon detection of the target molecules and cleavage or deactivation of the masking agent by the activated CasP6 protein.

[0255] In certain embodiments, the masking construct may suppress generation of a gene product. The gene product may be encoded by a reporter construct that is added to the sample. The masking construct may be an interfering RNA involved in a RNA interference pathway, such as a short hairpin RNA (shRNA) or small interfering RNA (siRNA). The masking construct may also comprise microRNA (miRNA). While present, the masking construct suppresses expression of the gene product. The gene product may be a fluorescent protein or other RNA transcript or proteins that would otherwise be detectable by a labeled probe, aptamer, or antibody but for the presence of the masking construct. Upon activation of the effector protein the masking construct is cleaved or otherwise silenced allowing for expression and detection of the gene product as the positive detectable signal.

[0256] In certain embodiments, the masking construct may sequester one or more reagents needed to generate a detectable positive signal such that release of the one or more reagents from the masking construct results in generation of the detectable positive signal. The one or more reagents may combine to produce a colorimetric signal, a chemiluminescent signal, a fluorescent signal, or any other detectable signal and may comprise any reagents known to be suitable for such purposes. In certain example embodiments, the one or more reagents are sequestered by RNA aptamers that bind the one or more reagents. The one or more reagents are released when the CasP6 protein is activated upon detection of a target molecule and the RNA aptamers are degraded.

[0257] In certain other embodiments, the masking construct binds to an immobilized reagent in solution thereby blocking the ability of the reagent to bind to a separate labeled binding partner that is free in solution. Thus, upon application of a washing step to a sample, the labeled binding partner can be washed out of the sample in the absence of a target molecule. However, if the effector protein is activated, the masking construct is cleaved to a degree sufficient to interfere with the ability of the masking construct to bind the reagent thereby allowing the labeled binding

partner to bind to the immobilized reagent. Thus, the labeled binding partner remains after the wash step indicating the presence of the target molecule in the sample. In certain aspects, the masking construct that binds the immobilized reagent is a RNA aptamer. The immobilized reagent may be a protein and the labeled binding partner may be a labeled antibody. Alternatively, the immobilized reagent may be streptavidin and the labeled binding partner may be labeled biotin. The label on the binding partner used in the above embodiments may be any detectable label known in the art. In addition, other known binding partners may be used in accordance with the overall design described herein.

[0258] In certain embodiments, the masking construct may comprise a ribozyme. Ribozymes are RNA molecules having catalytic properties. Ribozymes, both naturally and engineered, comprise or consist of RNA that may be targeted by the effector proteins disclosed herein. The ribozyme may be selected or engineered to catalyze a reaction that either generates a negative detectable signal or prevents generation of a positive control signal. Upon deactivation of the ribozyme by the activated effector protein the reaction generating a negative control signal, or preventing generation of a positive detectable signal, is removed thereby allowing a positive detectable signal to be generated. In one example embodiment, the ribozyme may catalyze a colorimetric reaction causing a solution to appear as a first color. When the ribozyme is deactivated the solution then turns to a second color, the second color being the detectable positive signal. An example of how ribozymes can be used to catalyze a colorimetric reaction are described in Zhao et al. "Signal amplification of glucosamine-6-phosphate based on ribozyme glmS," Biosens Bioelectron. 2014; 16:337-42, and provide an example of how such a system could be modified to work in the context of the embodiments disclosed herein. Alternatively, ribozymes, when present can generate cleavage products of, for example, RNA transcripts. Thus, detection of a positive detectable signal may comprise detection of non-cleaved RNA transcripts that are only generated in the absence of the ribozyme.

[0259] In certain embodiments, the one or more reagents is a protein, such as an enzyme, capable of facilitating generation of a detectable signal, such as a colorimetric, chemiluminescent, or fluorescent signal, that is inhibited or sequestered such that the protein cannot generate the detectable signal by the binding of one or more RNA aptamers to the protein. Upon activation of the effector proteins disclosed herein, the RNA aptamers are cleaved or degraded to an extent that they no longer inhibit the protein's ability to generate the detectable signal. In certain example embodiments, the aptamer is a thrombin inhibitor aptamer. In certain example embodiments the thrombin inhibitor aptamer has a sequence of GGGAACAAAGCUGAAGUACUUACCC (SEQ ID NO: 68). When this aptamer is cleaved, thrombin will become active and will cleave a peptide

colorimetric or fluorescent substrate. In certain example embodiments, the colorimetric substrate is para-nitroanilide (pNA) covalently linked to the peptide substrate for thrombin. Upon cleavage by thrombin, pNA is released and becomes yellow in color and easily visible to the eye. In certain example embodiments, the fluorescent substrate is 7-amino-4-methylcoumarin a blue fluorophore that can be detected using a fluorescence detector. Inhibitory aptamers may also be used for horseradish peroxidase (HRP), beta-galactosidase, or calf alkaline phosphatase (CAP) and within the general principals laid out above.

[0260] In certain embodiments, RNAse activity is detected colorimetrically via cleavage of enzyme-inhibiting aptamers. One potential mode of converting RNAse activity into a colorimetric signal is to couple the cleavage of an RNA aptamer with the re-activation of an enzyme that is capable of producing a colorimetric output. In the absence of RNA cleavage, the intact aptamer will bind to the enzyme target and inhibit its activity. The advantage of this readout system is that the enzyme provides an additional amplification step: once liberated from an aptamer via collateral activity (e.g. CasP6 collateral activity), the colorimetric enzyme will continue to produce colorimetric product, leading to a multiplication of signal.

[0261] In certain embodiments, an existing aptamer that inhibits an enzyme with a colorimetric readout is used. Several aptamer/enzyme pairs with colorimetric readouts exist, such as thrombin, protein C, neutrophil elastase, and subtilisin. These proteases have colorimetric substrates based upon pNA and are commercially available. In certain embodiments, a novel aptamer targeting a common colorimetric enzyme is used. Common and robust enzymes, such as beta-galactosidase, horseradish peroxidase, or calf intestinal alkaline phosphatase, could be targeted by engineered aptamers designed by selection strategies such as SELEX. Such strategies allow for quick selection of aptamers with nanomolar binding efficiencies and could be used for the development of additional enzyme/aptamer pairs for colorimetric readout.

[0262] In certain embodiments, RNAse activity is detected colorimetrically via cleavage of RNA-tethered inhibitors. Many common colorimetric enzymes have competitive, reversible inhibitors: for example, beta-galactosidase can be inhibited by galactose. Many of these inhibitors are weak, but their effect can be increased by increases in local concentration. By linking local concentration of inhibitors to RNAse activity, colorimetric enzyme and inhibitor pairs can be engineered into RNAse sensors. The colorimetric RNAse sensor based upon small-molecule inhibitors involves three components: the colorimetric enzyme, the inhibitor, and a bridging RNA that is covalently linked to both the inhibitor and enzyme, tethering the inhibitor to the enzyme. In the uncleaved configuration, the enzyme is inhibited by the increased local concentration of

the small molecule; when the RNA is cleaved (e.g. by CasP6 collateral cleavage), the inhibitor will be released and the colorimetric enzyme will be activated.

[0263] In certain embodiments, RNAse activity is detected colorimetrically via formation and/or activation of G-quadruplexes. G quadraplexes in DNA can complex with heme (iron (III)-protoporphyrin IX) to form a DNAzyme with peroxidase activity. When supplied with a peroxidase substrate (e.g. ABTS: (2,2'-Azinobis [3-ethylbenzothiazoline-6-sulfonic acid]-diammonium salt)), the G-quadraplex-heme complex in the presence of hydrogen peroxide causes oxidation of the substrate, which then forms a green color in solution. An example G-quadraplex forming DNA sequence is: GGGTAGGGCGGGTTGGGA (SEQ ID NO: 69). By hybridizing an RNA sequence to this DNA aptamer, formation of the G-quadraplex structure will be limited. Upon RNAse collateral activation (e.g. CasP6 collateral activation), the RNA staple will be cleaved allowing the G quadraplex to form and heme to bind. This strategy is particularly appealing because color formation is enzymatic, meaning there is additional amplification beyond RNAse activation.

[0264] In certain embodiments, the masking construct may be immobilized on a solid substrate in an individual discrete volume and sequesters a single reagent. For example, the reagent may be a bead comprising a dye. When sequestered by the immobilized reagent, the individual beads are too diffuse to generate a detectable signal, but upon release from the masking construct are able to generate a detectable signal, for example by aggregation or by a simple increase in solution concentration. In certain example embodiments, the immobilized masking agent is a RNA-based aptamer that can be cleaved by the activated effector protein upon detection of a target molecule. Individual discrete volumes are volumes separated from other volumes by space, a physical barrier, or any combination thereof. In certain embodiments, the individual discrete volume can be droplets (*e.g.*, a liquid drop of less than 1, 0.5, 0.2, 0.1, or 0.050 milliliters, a liquid drop of less than 45, 40, 20, 10, 5, 1, or 0.5 microliters, or a liquid drop of less than 450, 400, 200, 100, 50, 25, 10, 5, or 1 nanoliter); can be separated on a solid substrate, or can separated by microwells.

[0265] In one embodiment, the masking construct comprises a detection agent that changes color depending on whether the detection agent is aggregated or dispersed in solution. For example, certain nanoparticles, such as colloidal gold, undergo a visible purple to red color shift as they move from aggregates to dispersed particles. Accordingly, in certain embodiments, such detection agents may be held in aggregate by one or more bridge molecules. At least a portion of the bridge molecule comprises RNA. Upon activation of the effector proteins disclosed herein, the RNA portion of the bridge molecule is cleaved allowing the detection agent to disperse and

resulting in the corresponding change in color. In certain embodiments the, bridge molecule is a RNA molecule. In certain embodiments, the detection agent is a colloidal metal. The colloidal metal material may include water-insoluble metal particles or metallic compounds dispersed in a liquid, a hydrosol, or a metal sol. The colloidal metal may be selected from the metals in groups IA, IB, IIB and IIIB of the periodic table, as well as the transition metals, especially those of group VIII. Preferred metals include gold, silver, aluminum, ruthenium, zinc, iron, nickel and calcium. Other suitable metals also include the following in all of their various oxidation states: lithium, sodium, magnesium, potassium, scandium, titanium, vanadium, chromium, manganese, cobalt, copper, gallium, strontium, niobium, molybdenum, palladium, indium, tin, tungsten, rhenium, platinum, and gadolinium. The metals are preferably provided in ionic form, derived from an appropriate metal compound, for example the A13+, Ru3+, Zn2+, Fe3+, Ni2+ and Ca2+ ions.

[0266] When the RNA bridge is cut by the activated CasP6, the before mentioned color shift is observed. In certain embodiments the particles are colloidal metals. In certain other embodiments, the colloidal metal is a colloidal gold. In certain embodiments, the colloidal nanoparticles are 15 nm gold nanoparticles (AuNPs). Due to the unique surface properties of colloidal gold nanoparticles, maximal absorbance is observed at 520 nm when fully dispersed in solution and appear red in color to the naked eye. Upon aggregation of AuNPs, they exhibit a red-shift in maximal absorbance and appear darker in color, eventually precipitating from solution as a dark purple aggregate. In certain embodiments the nanoparticles are modified to include DNA linkers extending from the surface of the nanoparticle. Individual particles are linked together by single-stranded RNA (ssRNA) bridges that hybridize on each end of the RNA to at least a portion of the DNA linkers. Thus, the nanoparticles will form a web of linked particles and aggregate, appearing as a dark precipitate. Upon activation of the CRISPR effectors disclosed herein, the ssRNA bridge will be cleaved, releasing the AuNPs from the linked mesh and producing a visible red color. Thiol linkers on the end of the DNA linkers may be used for surface conjugation to the AuNPs. Other forms of conjugation may be used. In certain embodiments, two populations of AuNPs may be generated, one for each DNA linker. This will help facilitate proper binding of the ssRNA bridge with proper orientation. In certain embodiments, a first DNA linker is conjugated by the 3' end while a second DNA linker is conjugated by the 5' end.

[0267] In certain other embodiments, the masking construct may comprise an RNA oligonucleotide to which are attached a detectable label and a masking agent of that detectable label. An example of such a detectable label/masking agent pair is a fluorophore and a quencher of the fluorophore. Quenching of the fluorophore can occur as a result of the formation of a non-

fluorescent complex between the fluorophore and another fluorophore or non-fluorescent molecule. This mechanism is known as ground-state complex formation, static quenching, or contact quenching. Accordingly, the RNA oligonucleotide may be designed so that the fluorophore and quencher are in sufficient proximity for contact quenching to occur. Fluorophores and their cognate quenchers are known in the art and can be selected for this purpose by one having ordinary skill in the art. The particular fluorophore/quencher pair is not critical in the context of this invention, only that selection of the fluorophore/quencher pairs ensures masking of the fluorophore. Upon activation of the effector proteins disclosed herein, the RNA oligonucleotide is cleaved thereby severing the proximity between the fluorophore and quencher needed to maintain the contact quenching effect. Accordingly, detection of the fluorophore may be used to determine the presence of a target molecule in a sample. In certain other embodiments, the masking construct may comprise one or more RNA oligonucleotides to which are attached one or more metal nanoparticles, such as gold nanoparticles. In some embodiments, the masking construct comprises a plurality of metal nanoparticles crosslinked by a plurality of RNA oligonucleotides forming a closed loop. In one embodiment, the masking construct comprises three gold nanoparticles crosslinked by three RNA oligonucleotides forming a closed loop. In some embodiments, the cleavage of the RNA oligonucleotides by the CasP6 protein leads to a detectable signal produced by the metal nanoparticles.

[0268] In certain other embodiments, the masking construct may comprise one or more RNA oligonucleotides to which are attached one or more quantum dots. In some embodiments, the cleavage of the RNA oligonucleotides by the CasP6 protein leads to a detectable signal produced by the quantum dots.

[0269] In one embodiment, the masking construct may comprise a quantum dot. The quantum dot may have multiple linker molecules attached to the surface. At least a portion of the linker molecule comprises RNA. The linker molecule is attached to the quantum dot at one end and to one or more quenchers along the length or at terminal ends of the linker such that the quenchers are maintained in sufficient proximity for quenching of the quantum dot to occur. The linker may be branched. As above, the quantum dot/quencher pair is not critical, only that selection of the quantum dot/quencher pair ensures masking of the fluorophore. Quantum dots and their cognate quenchers are known in the art and can be selected for this purpose by one having ordinary skill in the art. Upon activation of the effector proteins disclosed herein, the RNA portion of the linker molecule is cleaved thereby eliminating the proximity between the quantum dot and one or more quenchers needed to maintain the quenching effect. In certain embodiments the quantum dot is streptavidin conjugated. RNA are attached via biotin linkers and recruit quenching molecules

with the sequences /5Biosg/UCUCGUACGUUC/3IAbRQSp/ (SEQ ID NO: 70) or /5Biosg/UCUCGUACGUUCUCUCGUACGUUC/3IAbRQSp/ (SEQ ID NO: 71), where /5Biosg/ is a biotin tag and /31AbRQSp/ is an Iowa Black® quencher. Upon cleavage by the activated effectors disclosed herein the quantum dot will fluoresce visibly.

[0270] In a similar fashion, fluorescence energy transfer (FRET) may be used to generate a detectable positive signal. FRET is a non-radiative process by which a photon from an energetically excited fluorophore (i.e. "donor fluorophore") raises the energy state of an electron in another molecule (i.e. "the acceptor") to higher vibrational levels of the excited singlet state. The donor fluorophore returns to the ground state without emitting a fluoresce characteristic of that fluorophore. The acceptor can be another fluorophore or non-fluorescent molecule. If the acceptor is a fluorophore, the transferred energy is emitted as fluorescence characteristic of that fluorophore. If the acceptor is a non-fluorescent molecule the absorbed energy is loss as heat. Thus, in the context of the embodiments disclosed herein, the fluorophore/quencher pair is replaced with a donor fluorophore/acceptor pair attached to the oligonucleotide molecule. When intact, the masking construct generates a first signal (negative detectable signal) as detected by the fluorescence or heat emitted from the acceptor. Upon activation of the effector proteins disclosed herein the RNA oligonucleotide is cleaved and FRET is disrupted such that fluorescence of the donor fluorophore is now detected (positive detectable signal).

[0271] In certain embodiments, the masking construct comprises the use of intercalating dyes which change their absorbance in response to cleavage of long RNAs to short nucleotides. Several such dyes exist. For example, pyronine-Y will complex with RNA and form a complex that has an absorbance at 572 nm. Cleavage of the RNA results in loss of absorbance and a color change. Methylene blue may be used in a similar fashion, with changes in absorbance at 688 nm upon RNA cleavage. Accordingly, in certain example embodiments the masking construct comprises a RNA and intercalating dye complex that changes absorbance upon the cleavage of RNA by the CasP6 proteins disclosed herein.

[0272] In certain embodiments, the masking construct may comprise an initiator for an HCR reaction. See e.g. Dirks and Pierce. PNAS 101, 15275-15728 (2004). HCR reactions utilize the potential energy in two hairpin species. When a single-stranded initiator having a portion of complementary to a corresponding region on one of the hairpins is released into the previously stable mixture, it opens a hairpin of one species. This process, in turn, exposes a single-stranded region that opens a hairpin of the other species. This process, in turn, exposes a single stranded region identical to the original initiator. The resulting chain reaction may lead to the formation of a nicked double helix that grows until the hairpin supply is exhausted. Detection of the resulting

products may be done on a gel or colorimetrically. Example colorimetric detection methods include, for example, those disclosed in Lu et al. "Ultra-sensitive colorimetric assay system based on the hybridization chain reaction-triggered enzyme cascade amplification ACS Appl Mater Interfaces, 2017, 9(1):167-175, Wang et al. "An enzyme-free colorimetric assay using hybridization chain reaction amplification and split aptamers" Analyst 2015, 150, 7657-7662, and Song et al. "Non-covalent fluorescent labeling of hairpin DNA probe coupled with hybridization chain reaction for sensitive DNA detection." Applied Spectroscopy, 70(4): 686-694 (2016).

[0273] In certain embodiments, the masking construct may comprise a HCR initiator sequence and a cleavable structural element, such as a loop or hairpin, that prevents the initiator from initiating the HCR reaction. Upon cleavage of the structure element by an activated CasP6 protein, the initiator is then released to trigger the HCR reaction, detection thereof indicating the presence of one or more targets in the sample. In certain embodiments, the masking construct comprises a hairpin with a RNA loop. When an activated CasP6 protein cuts the RNA loop, the initiator can be released to trigger the HCR reaction.

Embodiments

[0274] Various embodiments of the systems, methods, and compositions provided herein are included in the following non-limiting list of embodiments.

1. A composition comprising:

    a) a CasP6 polypeptide, or a nucleic acid molecule encoding the CasP6 polypeptide; and

    b) a CasP6 guide RNA, or one or more DNA molecules encoding the CasP6 guide RNA.

2. The composition of embodiment 1, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

3. The composition of embodiment 1 or embodiment 2, wherein the CasP6 guide RNA is a single guide RNA.

4. The composition of embodiment 3, wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

5. The composition of any one of embodiments 1-4, wherein the composition comprises a lipid.

6. The composition of any one of embodiments 1-4, wherein a) and b) are within a liposome.

7. The composition of any one of embodiments 1-4, wherein a) and b) are within a particle.

8. The composition of any one of embodiments 1-7, comprising one or more of: a buffer, a nuclease inhibitor, and a protease inhibitor.

9. The composition of any one of embodiments 1-8, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

10. The composition of any one of embodiments 1-9, wherein the CasP6 polypeptide is a nickase that can cleave only one strand of a double-stranded target nucleic acid molecule.

11. The composition of any one of embodiments 1-9, wherein the CasP6 polypeptide is a catalytically inactive CasP6 Polypeptide (dCasP6).

12. The composition of embodiment 10 or embodiment 11, wherein the CasP6 polypeptide comprises one or more mutations at a position corresponding to those selected from: D434 and E617 of SEQ ID NO: 1, D443 and E626 of SEQ ID NO: 2, or D503 and E750 of SEQ ID NO: 3.

13. The composition of any one of embodiments 1-12, further comprising a DNA donor template.

14. A CasP6 single guide RNA molecule, comprising:

a) a guide sequence that hybridizes to a target nucleic acid, and a duplex-forming segment; and

b) an activator sequence that can bind a CasP6 polypeptide.

15. The CasP6 single guide RNA molecule of embodiment 14, wherein the guide sequence has a length of from 20 to 33 nucleotides or wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

16. A DNA molecule comprising a nucleotide sequence encoding the CasP6 single guide RNA molecule of embodiment 14 or embodiment 15.

17. The DNA molecule of Embodiment 16, wherein the nucleotide sequence encoding the CasP6 single guide RNA is operably linked to a promoter.

18. The DNA molecule of Embodiment 17, wherein the promoter is functional in a eukaryotic cell.

19. The DNA molecule of Embodiment 18, wherein the promoter is functional in one or more of: a plant cell, a fungal cell, an animal cell, cell of an invertebrate, a fly cell, a cell of a vertebrate, a mammalian cell, a primate cell, a non-human primate cell, and a human cell.

20. The DNA molecule of any one of Embodiments 17-19, wherein the promoter is one or more of: a constitutive promoter, an inducible promoter, a cell type-specific promoter, and a tissue-specific promoter.

21. The DNA molecule of any one of Embodiments 16-20, wherein the DNA molecule is a recombinant expression vector.

22. The DNA molecule of Embodiment 21, wherein the recombinant expression vector is a recombinant adenoassociated viral vector, a recombinant retroviral vector, or a recombinant lentiviral vector.

23. The DNA molecule of Embodiment 17, wherein the promoter is functional in a prokaryotic cell.

24. A CasP6 fusion polypeptide comprising: a CasP6 polypeptide fused to a heterologous polypeptide.

25. The CasP6 fusion polypeptide of Embodiment 24, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

26. The CasP6 fusion polypeptide of Embodiment 24, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

27. The CasP6 fusion polypeptide of any one of embodiments 24-26, wherein the CasP6 polypeptide is a nickase that can cleave only one strand of a double-stranded target nucleic acid molecule.

28. The CasP6 fusion polypeptide of any one of embodiments 24-27, wherein the CasP6 polypeptide is a catalytically inactive CasP6 Polypeptide (dCasP6).

29. The CasP6 fusion polypeptide of embodiment 27 or embodiment 28, wherein the CasP6 polypeptide comprises one or more mutations at a position corresponding to those selected from: D434 and E617 of SEQ ID NO: 1, D443 and E626 of SEQ ID NO: 2, or D503 and E750 of SEQ ID NO: 3.

30. The CasP6 fusion polypeptide of any one of embodiments 24-29, wherein the heterologous polypeptide is fused to the N-terminus and/or the C-terminus of the CasP6 polypeptide.

31. The CasP6 fusion polypeptide of any one of embodiments 24-30, comprising a nuclear localization signal (NLS).

32. The CasP6 fusion polypeptide of any one of embodiments 24-31, wherein the heterologous polypeptide is a targeting polypeptide that provides for binding to a cell surface moiety on a target cell or target cell type.

33. The CasP6 fusion polypeptide of any one of embodiments 24-31, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies target DNA or RNA.

34. The CasP6 fusion polypeptide of embodiment 33, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, DNA repair activity, DNA damage activity, deamination activity,

dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity and glycosylase activity.

35. The CasP6 fusion polypeptide of embodiment 34, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, deamination activity, depurination activity, integrase activity, transposase activity, and recombinase activity.

36. The CasP6 fusion polypeptide of any one of embodiments 24-31, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies a target polypeptide associated with a target nucleic acid.

37. The CasP6 fusion polypeptide of embodiment 36, wherein the heterologous polypeptide exhibits histone modification activity.

38. The CasP6 fusion polypeptide of embodiment 36 or embodiment 37, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, deacetylase activity, kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity, deSUMOylating activity, ribosylation activity, deribosylation activity, myristoylation activity, demyristoylation activity, glycosylation activity (e.g., from O-GlcNAc transferase) and deglycosylation activity.

39. The CasP6 fusion polypeptide of embodiment 38, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, and deacetylase activity.

40. The CasP6 fusion polypeptide of any one of embodiments 24-31, wherein the heterologous polypeptide is an endosomal escape polypeptide.

41. The CasP6 fusion polypeptide of embodiment 40, wherein the endosomal escape polypeptide comprises an amino acid sequence set forth in SEQ ID NO: 31 or SEQ ID NO: 32..

42. The CasP6 fusion polypeptide of any one of embodiments 24-31, wherein the heterologous polypeptide is a chloroplast transit peptide.

43. The CasP6 fusion polypeptide of embodiment 42, wherein the chloroplast transit peptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

44. The CasP6 fusion polypeptide of any one of embodiments 24-31, wherein the heterologous polypeptide protein that increases or decreases transcription.

45. The CasP6 fusion polypeptide of embodiment 44, wherein the heterologous polypeptide is a transcriptional repressor domain.

46. The CasP6 fusion polypeptide of embodiment 44, wherein the heterologous polypeptide is a transcriptional activation domain.

47. The CasP6 fusion polypeptide of any one of embodiments 24-31, wherein the heterologous polypeptide comprises a protein binding domain.

48. A nucleic acid molecule encoding the CasP6 fusion polypeptide of any one of embodiments 24-47.

49. The nucleic acid molecule of Embodiment 48, wherein the nucleotide sequence encoding the CasP6 fusion polypeptide is operably linked to a promoter.

50. The nucleic acid molecule of Embodiment 49, wherein the promoter is functional in a eukaryotic cell.

51. The nucleic acid molecule of Embodiment 50, wherein the promoter is functional in one or more of: a plant cell, a fungal cell, an animal cell, cell of an invertebrate, a fly cell, a cell of a vertebrate, a mammalian cell, a primate cell, a non-human primate cell, and a human cell.

52. The nucleic acid molecule of any one of Embodiments 49-51, wherein the promoter is one or more of: constitutive promoter, an inducible promoter, a cell type-specific promoter, and a tissue-specific promoter.

53. The nucleic acid molecule of any one of Embodiments 48-52, wherein the DNA molecule is a recombinant expression vector.

54. The nucleic acid molecule of Embodiment 53, wherein the recombinant expression vector is a recombinant adenoassociated viral vector, a recombinant retroviral vector, or a recombinant lentiviral vector.

55. The nucleic acid molecule of Embodiment 49, wherein the promoter is functional in a prokaryotic cell.

56. The nucleic acid molecule of Embodiment 48, wherein the nucleic acid molecule is an mRNA.

57. One or more nucleic molecules encoding:

(a) a CasP6 guide RNA comprising an activator RNA and a guide RNA; and

(b) a CasP6 polypeptide.

58. The one or more nucleic acid molecules of embodiment 57, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

59. The one or more nucleic acid molecules of embodiment 57, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

60. The one or more nucleic acid molecules of any one of embodiments 57-59, wherein the CasP6 guide RNA is a single guide RNA.

61. The one or more nucleic acid molecules of any one of embodiments 57-59, wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

62. The one or more nucleic acid molecules of embodiment 61, wherein said one or more nucleic acid molecules comprises a first nucleotide sequence encoding the activator and a second nucleotide sequence encoding the guide, and wherein said first and second nucleotide sequences are present on different DNA molecules.

63. The one or more nucleic acid molecules of any one of embodiments 57-62, wherein said one or more nucleic acid molecules comprises a nucleotide sequence encoding the CasP6 polypeptide that is operably linked to a promoter.

64. The one or more nucleic acid molecules of Embodiment 63, wherein the promoter is functional in a eukaryotic cell.

65. The one or more nucleic acid molecules of Embodiment 64, wherein the promoter is functional in one or more of: a plant cell, a fungal cell, an animal cell, cell of an invertebrate, a fly cell, a cell of a vertebrate, a mammalian cell, a primate cell, a non-human primate cell, and a human cell.

66. The one or more nucleic acid molecules of any one of Embodiments 63-65, wherein the promoter is one or more of: a constitutive promoter, an inducible promoter, a cell type-specific promoter, and a tissue- specific promoter.

67. The one or more nucleic acid molecules of any one of Embodiments 57-66, wherein the one or more nucleic acid molecules is one or more recombinant expression vectors.

68. The one or more nucleic acid molecules of Embodiment 67, wherein the one or more recombinant expression vectors are selected from: one or more adenoassociated viral vectors, one or more recombinant retroviral vectors, or one or more recombinant lenti viral vectors.

69. The one or more nucleic acid molecules of Embodiment 63, wherein the promoter is functional in a prokaryotic cell.

70. A eukaryotic cell comprising one or more of:

a) a CasP6 polypeptide, or a nucleic acid molecule encoding the CasP6 polypeptide,

b) a CasP6 guide RNA, or a nucleic acid molecule encoding the CasP6 guide RNA.

71. The eukaryotic cell of embodiment 70, comprising the nucleic acid molecule encoding the CasP6 polypeptide, wherein said nucleic acid molecule is integrated into the genomic DNA of the cell.

72. The eukaryotic cell of embodiment 70 or embodiment 71, wherein the eukaryotic cell is a plant cell, a mammalian cell, an insect cell, an arachnid cell, a fungal cell, a bird cell, a reptile cell, an amphibian cell, an invertebrate cell, a mouse cell, a rat cell, a primate cell, a non-human primate cell, or a human cell.

73. A cell comprising a CasP6 fusion polypeptide, or a nucleic acid molecule encoding the CasP6 fusion polypeptide.

74. The cell of embodiment 73, wherein the cell is a prokaryotic cell.

75. The cell of embodiment 73 or embodiment 74, comprising the nucleic acid molecule encoding the CasP6 fusion polypeptide, wherein said nucleic acid molecule is integrated into the genomic DNA of the cell.

76. A method of modifying a target nucleic acid, the method comprising contacting the target nucleic acid with:

a) a CasP6 polypeptide; and

b) a CasP6 guide RNA comprising a guide sequence that hybridizes to a target sequence of the target nucleic acid,

wherein said contacting results in modification of the target nucleic acid by the CasP6 polypeptide.

77. The method of embodiment 76, wherein said modification is cleavage of the target nucleic acid.

78. The method of embodiment 76 or embodiment 77, wherein the target nucleic acid is selected from: double stranded DNA, single stranded DNA, RNA, genomic DNA, and extrachromosomal DNA.

79. The method of any of embodiments 76-78, wherein said contacting takes place in vitro outside of a cell.

80. The method of any of embodiments 76-78, wherein said contacting takes place inside of a cell in culture.

81. The method of any of embodiments 76-78, wherein said contacting takes place inside of a cell in vivo.

82. The method of embodiment 80 or embodiment 81, wherein the cell is a eukaryotic cell.

83. The method of embodiment 82, wherein the cell is selected from: a plant cell, a fungal cell, a mammalian cell, a reptile cell, an insect cell, an avian cell, a fish cell, a parasite cell, an arthropod

cell, a cell of an invertebrate, a cell of a vertebrate, a rodent cell, a mouse cell, a rat cell, a primate cell, a non- human primate cell, and a human cell.

84. The method of embodiment 80 or embodiment 81, wherein the cell is a prokaryotic cell.

85. The method of any one of embodiments 76-84, wherein said contacting results in genome editing or cleavage of a target RNA.

86. The method of any one of embodiments 76-78 and 80-85, wherein said contacting comprises: introducing into a cell: (a) the CasP6 polypeptide, or a nucleic acid molecule encoding the CasP6 polypeptide, and (b) the CasP6 guide RNA, or a nucleic acid molecule encoding the CasP6 guide RNA.

87. The method of embodiment 86, wherein said contacting further comprises: introducing a DNA donor template into the cell.

88. The method of any one of embodiments 76-87, wherein the CasP6 guide RNA is a single guide RNA.

89. The method of any one of embodiments 88, wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

90. A method of modulating transcription from a target DNA, modifying a target nucleic acid, or modifying a protein associated with a target nucleic acid, the method comprising contacting the target nucleic acid with:

a) a CasP6 fusion polypeptide comprising a CasP6 polypeptide fused to a heterologous polypeptide; and

b) a CasP6 guide RNA comprising a guide sequence that hybridizes to a target sequence of the target nucleic acid.

91. The method of embodiment 90, wherein the CasP6 guide RNA is a single guide RNA.

92. The method of embodiment 91, wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

93. The method of any of embodiments 90-92, wherein said modification is not cleavage of the target nucleic acid.

94. The method of any of embodiments 90-93, wherein the target nucleic acid is selected from: double stranded DNA, single stranded DNA, RNA, genomic DNA, and extrachromosomal DNA.

95. The method of any of embodiments 90-94, wherein said contacting takes place in vitro outside of a cell.

96. The method of any of embodiments 90-94, wherein said contacting takes place inside of a cell in culture.

97. The method of any of embodiments 90-94, wherein said contacting takes place inside of a cell in vivo.

98. The method of embodiment 96 or embodiment 97, wherein the cell is a eukaryotic cell.

99. The method of embodiment 98, wherein the cell is selected from: a plant cell, a fungal cell, a mammalian cell, a reptile cell, an insect cell, an avian cell, a fish cell, a parasite cell, an arthropod cell, a cell of an invertebrate, a cell of a vertebrate, a rodent cell, a mouse cell, a rat cell, a primate cell, a non- human primate cell, and a human cell.

100. The method of embodiment 96 or embodiment 97, wherein the cell is a prokaryotic cell.

101. The method of any one of embodiments 90-100, wherein said contacting comprises: introducing into a cell: (a) the CasP6 fusion polypeptide, or a nucleic acid molecule encoding the CasP6 fusion polypeptide, and (b) the CasP6 guide RNA, or a nucleic acid molecule encoding the CasP6 guide RNA.

102. The method of any one of embodiments 90-101, wherein the CasP6 polypeptide is a catalytically inactive CasP6 Polypeptide (dCasP6).

103. The method of any one of embodiments 90-102, wherein the CasP6 polypeptide comprises one or more mutations at a position corresponding to those selected from: D434 and E617 of SEQ ID NO: 1, D443 and E626 of SEQ ID NO: 2, or D503 and E750 of SEQ ID NO: 3.

104. The method of any one of embodiments 90-103, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies target DNA.

105. The method of embodiment 104, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, DNA repair activity, DNA damage activity, deamination activity, dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity and glycosylase activity.

106. The method of embodiment 105, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, deamination activity, depurination activity, integrase activity, transposase activity, and recombinase activity.

107. The method of any one of embodiments 90-103, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies a target polypeptide associated with a target nucleic acid.

108. The method of embodiment 107, wherein the heterologous polypeptide exhibits histone modification activity.

109. The method of embodiment 107 or embodiment 108, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, deacetylase activity, kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity, deSUMOylating activity, ribosylation activity, deribosylation activity, myristoylation activity, demyristoylation activity, glycosylation activity (e.g., from O-GlcNAc transferase) and deglycosylation activity.

110. The method of embodiment 109, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, and deacetylase activity.

111. The method of any one of embodiments 90-103, wherein the heterologous polypeptide is protein that increases or decreases transcription.

112. The method of embodiment 111, wherein the heterologous polypeptide is a transcriptional repressor domain.

113. The method of embodiment 111, wherein the heterologous polypeptide is a transcriptional activation domain.

114. The method of any one of embodiments 90-103, wherein the heterologous polypeptide is a protein binding domain.

115. A transgenic, multicellular, non-human organism whose genome comprises a transgene comprising a nucleotide sequence encoding one or more of:

a) a CasP6 polypeptide, and/or

b) a CasP6 guide RNA.

116. The transgenic, multicellular, non-human organism of embodiment 115, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more amino acid sequence identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

117. The transgenic, multicellular, non-human organism of embodiment 115, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more amino acid sequence identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

118. The transgenic, multicellular, non-human organism of any one of embodiments 115-117, wherein the organism is a plant, a monocotyledon plant, a dicotyledon plant, an invertebrate animal, an insect, an arthropod, an arachnid, a parasite, a worm, a cnidarian, a vertebrate animal,

a fish, a reptile, an amphibian, an ungulate, a bird, a pig, a horse, a sheep, a rodent, a mouse, a rat, or a non-human primate.

119. A system comprising:

a) a CasP6 polypeptide and a CasP6 single guide RNA;

b) a CasP6 polypeptide, a CasP6 guide RNA, and a DNA donor template;

c) a CasP6 fusion polypeptide and a CasP6 guide RNA;

d) a CasP6 fusion polypeptide, a CasP6 guide RNA, and a DNA donor template;

e) an mRNA encoding a CasP6 polypeptide, and a CasP6 single guide RNA;

f) an mRNA encoding a CasP6 polypeptide; a CasP6 guide RNA, and a DNA donor template;

g) an mRNA encoding a CasP6 fusion polypeptide, and a CasP6 guide RNA;

h) an mRNA encoding a CasP6 fusion polypeptide, a CasP6 guide RNA, and a DNA donor template;

i) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasP6 polypeptide; and ii) a nucleotide sequence encoding a CasP6 guide RNA;

j) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasP6 polypeptide; ii) a nucleotide sequence encoding a CasP6 guide RNA; and iii) a DNA donor template;

k) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasP6 fusion polypeptide; and ii) a nucleotide sequence encoding a CasP6 guide RNA; or

l) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasP6 fusion polypeptide; ii) a nucleotide sequence encoding a CasP6 guide RNA; and a DNA donor template.

120. The CasP6 system of embodiment 119, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more amino acid sequence identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

121. The CasP6 system of embodiment 119, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more amino acid sequence identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

122. The composition of embodiments 13, wherein the DNA donor template has a length of from 8 nucleotides to 1000 nucleotides.

123. The composition of embodiment 13, wherein the DNA donor template has a length of from 25 nucleotides to 500 nucleotides.

124. The method of embodiment 87, wherein the DNA donor template has a length of from 8 nucleotides to 1000 nucleotides.

125. The method of embodiment 87, wherein the DNA donor template has a length of from 25 nucleotides to 500 nucleotides.

126. A nucleic acid detection system comprising: (i) a CasP6 polypeptide set forth in any one of embodiments 1, 2, or 9; (ii) one or more CasP6 guide RNAs each comprising a guide sequence that can hybridize to RNA molecules having corresponding target sequences; and (iii) an RNA-based masking construct.

127. The detection system of embodiment 126, wherein the RNA-based masking construct does not comprise the corresponding target sequence(s).

128. A method of detecting a target RNA molecule comprising:

(a) contacting target RNA molecule with: (i) a CasP6 polypeptide of embodiment 1, 2, or 9; or a CasP6 fusion polypeptide of any one of embodiments 24-26, 30-46, or 47; (ii) one or more CasP6 guide RNAs each comprising a guide sequence that can hybridize to the target RNA molecule; and (iii) an RNA-based masking construct, wherein the CasP6 comprises endonuclease enzyme activity, and wherein the RNA-based masking construct does not comprise the target sequence; and

(b) assaying for modification of the RNA-based masking construct.

Examples

[0275] The following examples are not intended to limit the scope of what the inventors regard as their invention.

Example 1 -- Nuclease effector sequence

[0276] Sequence date from samples of microbial communities is analyzed to identify new Class 2 CRISPR-Cas systems. Candidate sequences are proposed based on proximity to CRISPR arrays and the presence of conserved sequence domains.

[0277] Figure 1 shows of nuclease sequences here termed CasP6. As diagramed in Figure 2, CasP6 polypeptides contain a split RuvC domain in the C-terminal region (RuvC-I, RuvC-II, and RuvC-III).

Example 2 -- crRNA

[0278] A CRISPR array adjacent to the nuclease of Example 1 indicates functional RNA components. The direct repeat (DR) sequences of the CRISPR array are shown in Figure 3 (SEQ ID NOs: 4-6).

Example 3 -- Single guide sequence

[0279] A vector is made having an Arabidopis U6 (Pol III) promoter is driving expression of an RNA made up of the 3' terminal 21 residues of RNA encoded by the SEQ ID NO: 4 sequence

(AGTGCATCGAATAGATGCAAC; SEQ ID NO: 72) fused to a 22 residue guide RNA that can hybridize to the Phytoene desaturase (PDS) gene of soybean.

Example 4 – CasP6

[0280] An effector nuclease vector is made by making a vector having a 35S promoter operably linked to drive expression of a soybean codon-optimized mRNA, having an amino acid sequence of CasP6A flanked at both the N and C termini by nuclear localization sequences.

Example 5 -- Genome editing

[0281] The vectors of examples 3 and 4 are co-transformed into a soybean plant. Transformed plants are regenerated, and the albino phenotype of PDS mutants is observed. The PDS genomic sequence of plants material with albino phenotype is sequenced. Mutations in the PDS gene are found.

Example 6 -- Ribonucleoparticles

[0282] Two RNAs are expressed in E coli.

[0283] The first RNA is translated into a polypeptide of CasP6B, SEQ ID NO: 2.

[0284] The second RNA comprises the sequence crRNA sequence encoded by AGTGCACCGCATAGGTGCAAC (3' terminus of SEQ ID NO 5; SEQ ID NO: 73), followed at the 3' end by the spacer sequence aactcgtaattcacagttca. This sequence is complementary to the human blue eye color single nucleotide polymorphism.

[0285] The polypeptide bound to the two RNA molecules is purified.

Example 7 – Diagnostic test

[0286] The sequences of the HERC2 gene (determinative of eye color) are amplified by PCR from different human genomic DNA sample, using a T7 exonuclease resistant (phosphorothioate-containing) primer for coding strand amplification. The complementary strands are then digested with T7 exonuclease, leaving single stranded HERC2 DNAs.

[0287] The ribonucleoparticles purified as in Example 6 are added to the single stranded HERC2 DNAs. The DNAs from genomic samples of blue eye colored individuals are digested by the ribonucloparticles. The DNAs from genomic samples of brown eye colored individuals are not digested by the ribonucloparticles.

Example 8 – RNA target

[0288] A polypeptide comprising CasP6B is produced in E. coli and purified.

[0289] A single stranded RNA substrate of about 300 nucleotides in length is produced by RNA from a PFSexpression library. The DNA encoding RNA targets has in part the following sequence: GGCgagtccgagcagaagaagaaNNNNNNNNGTCGAGTGCAAAACCTTTCGCGGTAT

GG (SEQ ID NO: 74). In this sequence, the Ns represent nucleotides that encode the randomized PFS candidate regions in the RNA substrate.

[0290] Three guide RNAs are made. The guides have at the 5' terminus the DR sequence ACAAGUGCACCGCAUAGGUGCAAC (SEQ ID NO: 75), fused at the 3' end to individual spacers. A first 16 nt spacer sequence is UUCUUCUUCUGCUCGG (SEQ ID NO: 76; guide 1, downstream of the PAM sites guide 1). A second 17nt spacer sequence is AAGGUUUUGCACUCGAC (SEQ ID NO: 77; guide 2, upstream of the PAM sites). A third spacer is the reverse complement of the 17 nt sequence of guide 2, to serve as a control in addition to controls lacking the polypeptide or any guide RNA.

[0291] The polypeptide and individual guide RNAs are mixed in NEB2 buffer at a 1:1 ratio, along with the single stranded RNA substrate, and incubated for different lengths of time. Unlike the control reactions, mixtures containing the polypeptide and either guide 1 or guide 2 produce detectable RNA fragments of amounts increasing with longer reaction time.

Example 9 – RNA target cleavage temperature optima

[0292] A polypeptide comprising CasP6B (SEQ ID NO: 2) is produced in E. coli and purified. Its RNA substrate cutting is tested in reaction mixtures at various temperatures. The results are shown in figures 4A and B.

[0293] As can be seen in figure 4A, CasP6B optimal targeted cutting occurs at the highest temperature on tested (42 Celsius), but the enzyme can also cut at physiological temperatures (27-37 Celsius).

[0294] The targeted ssRNA of figure 4B is fluorescently labeled, and the image shows only the fluorescence of the gel. Hence, crRNAs bands are not visible as in figure 4A. The optimal cutting temperature is in the 42-47 degrees Celsius range, with limited cutting at temperatures above 52 Celsius.

Example 10 –Spacer Length of Guide RNA

A polypeptide comprising CasP6B (SEQ ID NO: 2) is produced in E. coli and purified. Its RNA substrate cutting is tested in reaction mixtures at various temperatures. Four different guide RNAs are made with spacer lengths of 18, 19, 20, or 21 residues. The spacer is the targeting segment of the guide RNA. The guide with a spacer length of 18 residues cuts single stranded RNA most efficiently.

[0295] Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, certain changes and modifications may be practiced within the scope of the appended claims.

Claims

What is claimed is:

1. A composition comprising:

a) a CasP6 polypeptide, or a nucleic acid molecule encoding the CasP6 polypeptide; and

b) a CasP6 guide RNA, or one or more DNA molecules encoding the CasP6 guide RNA.

2. The composition of claim 1, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

3. The composition of claim 1, wherein the CasP6 guide RNA is a single guide RNA.

4. The composition of claim 3, wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

5. The composition of claim 1, wherein the composition comprises a lipid.

6. The composition of claim 1, wherein a) and b) are within a liposome.

7. The composition of claim 1, wherein a) and b) are within a particle.

8. The composition of claim 1, comprising one or more of: a buffer, a nuclease inhibitor, and a protease inhibitor.

9. The composition of claim 1, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

10. The composition of claim 1, wherein the CasP6 polypeptide is a nickase that can cleave only one strand of a double-stranded target nucleic acid molecule.

11. The composition of claim 1, wherein the CasP6 polypeptide is a catalytically inactive CasP6 Polypeptide (dCasP6).

12. The composition of claim 10 or claim 11, wherein the CasP6 polypeptide comprises one or more mutations at a position corresponding to those selected from: D434 and E617 of SEQ ID NO: 1, D443 and E626 of SEQ ID NO: 2, or D503 and E750 of SEQ ID NO: 3.

13. The composition of any one of claims 1, 10, or 11, further comprising a DNA donor template.

14. A CasP6 single guide RNA molecule, comprising:

a) a guide sequence that hybridizes to a target nucleic acid, and a duplex-forming segment; and

b) an activator sequence that can bind a CasP6 polypeptide.

15. The CasP6 single guide RNA molecule of claim 14, wherein the guide sequence has a length of from 20 to 33 nucleotides or wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

16. A DNA molecule comprising a nucleotide sequence encoding the CasP6 single guide RNA molecule of claim 14 or claim 15.

17. The DNA molecule of Claim 16, wherein the nucleotide sequence encoding the CasP6 single guide RNA is operably linked to a promoter.

18. The DNA molecule of Claim 17, wherein the promoter is functional in a eukaryotic cell.

19. The DNA molecule of Claim 18, wherein the promoter is functional in one or more of: a plant cell, a fungal cell, an animal cell, cell of an invertebrate, a fly cell, a cell of a vertebrate, a mammalian cell, a primate cell, a non-human primate cell, and a human cell.

20. The DNA molecule of any one of Claims 17-19, wherein the promoter is one or more of: a constitutive promoter, an inducible promoter, a cell type-specific promoter, and a tissue-specific promoter.

21. The DNA molecule of any one of Claims 16-19, wherein the DNA molecule is a recombinant expression vector.

22. The DNA molecule of Claim 21, wherein the recombinant expression vector is a recombinant adenoassociated viral vector, a recombinant retroviral vector, or a recombinant lentiviral vector.

23. The DNA molecule of Claim 17, wherein the promoter is functional in a prokaryotic cell.

24. A CasP6 fusion polypeptide comprising: a CasP6 polypeptide fused to a heterologous polypeptide.

25. The CasP6 fusion polypeptide of Claim 24, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

26. The CasP6 fusion polypeptide of Claim 24, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

27. The CasP6 fusion polypeptide of claim 24, wherein the CasP6 polypeptide is a nickase that can cleave only one strand of a double-stranded target nucleic acid molecule.

28. The CasP6 fusion polypeptide of claim 24, wherein the CasP6 polypeptide is a catalytically inactive CasP6 Polypeptide (dCasP6).

29. The CasP6 fusion polypeptide of claim 27 or claim 28, wherein the CasP6 polypeptide comprises one or more mutations at a position corresponding to those selected from: D434 and E617 of SEQ ID NO: 1, D443 and E626 of SEQ ID NO: 2, or D503 and E750 of SEQ ID NO: 3.

30. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, wherein the heterologous polypeptide is fused to the N-terminus and/or the C-terminus of the CasP6 polypeptide.

31. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, comprising a nuclear localization signal (NLS).

32. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, wherein the heterologous polypeptide is a targeting polypeptide that provides for binding to a cell surface moiety on a target cell or target cell type.

33. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies target DNA or RNA.

34. The CasP6 fusion polypeptide of claim 33, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, DNA repair activity, DNA damage activity, deamination activity, dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity and glycosylase activity.

35. The CasP6 fusion polypeptide of claim 34, wherein the heterologous polypeptide exhibits an one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, deamination activity, depurination activity, integrase activity, transposase activity, and recombinase activity.

36. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies a target polypeptide associated with a target nucleic acid.

37. The CasP6 fusion polypeptide of claim 36, wherein the heterologous polypeptide exhibits histone modification activity.

38. The CasP6 fusion polypeptide of claim 36, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, deacetylase activity, kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity, deSUMOylating activity, ribosylation activity, deribosylation activity, myristoylation activity, demyristoylation activity, glycosylation activity (e.g., from O-GlcNAc transferase) and deglycosylation activity.

39. The CasP6 fusion polypeptide of claim 38, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, and deacetylase activity.

40. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, wherein the heterologous polypeptide is an endosomal escape polypeptide.

41. The CasP6 fusion polypeptide of claim 40, wherein the endosomal escape polypeptide comprises an amino acid sequence set forth in SEQ ID NO: 31 or SEQ ID NO: 32..

42. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, wherein the heterologous polypeptide is a chloroplast transit peptide.

43. The CasP6 fusion polypeptide of claim 42, wherein the chloroplast transit peptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, SEQ ID NO: 24, SEQ ID NO: 25, SEQ ID NO: 26, SEQ ID NO: 27, SEQ ID NO: 28, SEQ ID NO: 29, and SEQ ID NO: 30.

44. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, wherein the heterologous polypeptide protein that increases or decreases transcription.

45. The CasP6 fusion polypeptide of claim 44, wherein the heterologous polypeptide is a transcriptional repressor domain.

46. The CasP6 fusion polypeptide of claim 44, wherein the heterologous polypeptide is a transcriptional activation domain.

47. The CasP6 fusion polypeptide of any one of claims 24, 27, or 28, wherein the heterologous polypeptide comprises a protein binding domain.

48. A nucleic acid molecule encoding the CasP6 fusion polypeptide of any one of claims 24, 27, or 28 .

49. The nucleic acid molecule of Claim 48, wherein the nucleotide sequence encoding the CasP6 fusion polypeptide is operably linked to a promoter.

50. The nucleic acid molecule of Claim 49, wherein the promoter is functional in a eukaryotic cell.

51. The nucleic acid molecule of Claim 50, wherein the promoter is functional in one or more of: a plant cell, a fungal cell, an animal cell, cell of an invertebrate, a fly cell, a cell of a vertebrate, a mammalian cell, a primate cell, a non-human primate cell, and a human cell.

52. The nucleic acid molecule of any one of Claims 49-51, wherein the promoter is one or more of: constitutive promoter, an inducible promoter, a cell type-specific promoter, and a tissue-specific promoter.

53. The nucleic acid molecule of claim 48, wherein the DNA molecule is a recombinant expression vector.

54. The nucleic acid molecule of Claim 53, wherein the recombinant expression vector is a recombinant adenoassociated viral vector, a recombinant retroviral vector, or a recombinant lentiviral vector.

55. The nucleic acid molecule of Claim 49, wherein the promoter is functional in a prokaryotic cell.

56. The nucleic acid molecule of Claim 48, wherein the nucleic acid molecule is an mRNA.

57. One or more nucleic molecules encoding:

(a) a CasP6 guide RNA comprising an activator RNA and a guide RNA; and

(b) a CasP6 polypeptide.

58. The one or more nucleic acid molecules of claim 57, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

59. The one or more nucleic acid molecules of claim 57, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

60. The one or more nucleic acid molecules of any one of claims 57-59, wherein the CasP6 guide RNA is a single guide RNA.

61. The one or more nucleic acid molecules of any one of claims 57-59, wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

62. The one or more nucleic acid molecules of claim 61, wherein said one or more nucleic acid molecules comprises a first nucleotide sequence encoding the activator and a second nucleotide sequence encoding the guide, and wherein said first and second nucleotide sequences are present on different DNA molecules.

63. The one or more nucleic acid molecules of any one of claims 57-59, wherein said one or more nucleic acid molecules comprises a nucleotide sequence encoding the CasP6 polypeptide that is operably linked to a promoter.

64. The one or more nucleic acid molecules of Claim 63, wherein the promoter is functional in a eukaryotic cell.

65. The one or more nucleic acid molecules of Claim 64, wherein the promoter is functional in one or more of: a plant cell, a fungal cell, an animal cell, cell of an invertebrate, a fly cell, a cell of a vertebrate, a mammalian cell, a primate cell, a non-human primate cell, and a human cell.

66. The one or more nucleic acid molecules of Claim 63, wherein the promoter is one or more of: a constitutive promoter, an inducible promoter, a cell type-specific promoter, and a tissue-specific promoter.

67. The one or more nucleic acid molecules of any one of Claims 57-59, wherein the one or more nucleic acid molecules is one or more recombinant expression vectors.

68. The one or more nucleic acid molecules of Claim 67, wherein the one or more recombinant expression vectors are selected from: one or more adenoassociated viral vectors, one or more recombinant retroviral vectors, or one or more recombinant lenti viral vectors.

69. The one or more nucleic acid molecules of Claim 63, wherein the promoter is functional in a prokaryotic cell.

70. A eukaryotic cell comprising one or more of:

a) a CasP6 polypeptide, or a nucleic acid molecule encoding the CasP6 polypeptide,

b) a CasP6 guide RNA, or a nucleic acid molecule encoding the CasP6 guide RNA.

71. The eukaryotic cell of claim 70, comprising the nucleic acid molecule encoding the CasP6 polypeptide, wherein said nucleic acid molecule is integrated into the genomic DNA of the cell.

72. The eukaryotic cell of claim 70 or claim 71, wherein the eukaryotic cell is a plant cell, a mammalian cell, an insect cell, an arachnid cell, a fungal cell, a bird cell, a reptile cell, an amphibian cell, an invertebrate cell, a mouse cell, a rat cell, a primate cell, a non-human primate cell, or a human cell.

73. A cell comprising a CasP6 fusion polypeptide, or a nucleic acid molecule encoding the CasP6 fusion polypeptide.

74. The cell of claim 73, wherein the cell is a prokaryotic cell.

75. The cell of claim 73 or claim 74, comprising the nucleic acid molecule encoding the CasP6 fusion polypeptide, wherein said nucleic acid molecule is integrated into the genomic DNA of the cell.

76. A method of modifying a target nucleic acid, the method comprising contacting the target nucleic acid with:

a) a CasP6 polypeptide; and

b) a CasP6 guide RNA comprising a guide sequence that hybridizes to a target sequence of the target nucleic acid,

wherein said contacting results in modification of the target nucleic acid by the CasP6 polypeptide.

77. The method of claim 76, wherein said modification is cleavage of the target nucleic acid.

78. The method of claim 76 or claim 77, wherein the target nucleic acid is selected from: double stranded DNA, single stranded DNA, RNA, genomic DNA, and extrachromosomal DNA.

79. The method of any of claims 76 or 77, wherein said contacting takes place in vitro outside of a cell.

80. The method of claim 76, wherein said contacting takes place inside of a cell in culture.

81. The method of claim 76, wherein said contacting takes place inside of a cell in vivo.

82. The method of claim 80 or claim 81, wherein the cell is a eukaryotic cell.

83. The method of claim 82, wherein the cell is selected from: a plant cell, a fungal cell, a mammalian cell, a reptile cell, an insect cell, an avian cell, a fish cell, a parasite cell, an arthropod cell, a cell of an invertebrate, a cell of a vertebrate, a rodent cell, a mouse cell, a rat cell, a primate cell, a non- human primate cell, and a human cell.

84. The method of claim 80 or claim 81, wherein the cell is a prokaryotic cell.

85. The method of any one of claims 76, 77, 80, or 81, wherein said contacting results in genome editing or cleavage of a target RNA.

86. The method of any one of claims 76, 77, 80, or 81, wherein said contacting comprises: introducing into a cell: (a) the CasP6 polypeptide, or a nucleic acid molecule encoding the CasP6 polypeptide, and (b) the CasP6 guide RNA, or a nucleic acid molecule encoding the CasP6 guide RNA.

87. The method of claim 86, wherein said contacting further comprises: introducing a DNA donor template into the cell.

88. The method of any one of claims 76, 77, 80, or 81, wherein the CasP6 guide RNA is a single guide RNA.

89. The method of claim 88, wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

90. A method of modulating transcription from a target DNA, modifying a target nucleic acid, or modifying a protein associated with a target nucleic acid, the method comprising contacting the target nucleic acid with:

a) a CasP6 fusion polypeptide comprising a CasP6 polypeptide fused to a heterologous polypeptide; and

b) a CasP6 guide RNA comprising a guide sequence that hybridizes to a target sequence of the target nucleic acid.

91. The method of claim 90, wherein the CasP6 guide RNA is a single guide RNA.

92. The method of claim 91, wherein the guide RNA comprises an RNA encoded by residues 17-37 of SEQ ID NO: 4, residues 18-38 of SEQ ID NO: 5, or residues 15-36 of SEQ ID NO: 6.

93. The method of any of claims 90-92, wherein said modification is not cleavage of the target nucleic acid.

94. The method of any of claims 90-92, wherein the target nucleic acid is selected from: double stranded DNA, single stranded DNA, RNA, genomic DNA, and extrachromosomal DNA.

95. The method of any of claims 90-92, wherein said contacting takes place in vitro outside of a cell.

96. The method of claim 90, wherein said contacting takes place inside of a cell in culture.

97. The method of claim 90, wherein said contacting takes place inside of a cell in vivo.

98. The method of claim 96 or claim 97, wherein the cell is a eukaryotic cell.

99. The method of claim 98, wherein the cell is selected from: a plant cell, a fungal cell, a mammalian cell, a reptile cell, an insect cell, an avian cell, a fish cell, a parasite cell, an arthropod cell, a cell of an invertebrate, a cell of a vertebrate, a rodent cell, a mouse cell, a rat cell, a primate cell, a non- human primate cell, and a human cell.

100. The method of claim 96 or claim 97, wherein the cell is a prokaryotic cell.

101. The method of any one of claims 90, 91, 92, 96, or 97, wherein said contacting comprises: introducing into a cell: (a) the CasP6 fusion polypeptide, or a nucleic acid molecule encoding the CasP6 fusion polypeptide, and (b) the CasP6 guide RNA, or a nucleic acid molecule encoding the CasP6 guide RNA.

102. The method of any one of claims 90, 91, 92, 96, or 97, wherein the CasP6 polypeptide is a catalytically inactive CasP6 Polypeptide (dCasP6).

103. The method of any one of claims 90, 91, 92, 96, or 97, wherein the CasP6 polypeptide comprises one or more mutations at a position corresponding to those selected from: D434 and E617 of SEQ ID NO: 1, D443 and E626 of SEQ ID NO: 2, or D503 and E750 of SEQ ID NO: 3.

104. The method of any one of claims 90, 91, 92, 96, or 97, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies target DNA.

105. The method of claim 104, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase activity, DNA repair activity, DNA damage activity, deamination activity, dismutase activity, alkylation activity, depurination activity, oxidation activity, pyrimidine dimer forming activity, integrase activity, transposase activity, recombinase activity, polymerase activity, ligase activity, helicase activity, photolyase activity and glycosylase activity.

106. The method of claim 105, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: nuclease activity, methyltransferase activity, demethylase

activity, deamination activity, depurination activity, integrase activity, transposase activity, and recombinase activity.

107. The method of any one of claims 90, 91, 92, 96, or 97, wherein the heterologous polypeptide exhibits an enzymatic activity that modifies a target polypeptide associated with a target nucleic acid.

108. The method of claim 107, wherein the heterologous polypeptide exhibits histone modification activity.

109. The method of claim 107, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, deacetylase activity, kinase activity, phosphatase activity, ubiquitin ligase activity, deubiquitinating activity, adenylation activity, deadenylation activity, SUMOylating activity, deSUMOylating activity, ribosylation activity, deribosylation activity, myristoylation activity, demyristoylation activity, glycosylation activity (e.g., from O-GlcNAc transferase) and deglycosylation activity.

110. The method of claim 109, wherein the heterologous polypeptide exhibits one or more enzymatic activities selected from: methyltransferase activity, demethylase activity, acetyltransferase activity, and deacetylase activity.

111. The method of any one of claims 90, 91, 92, 96, or 97, wherein the heterologous polypeptide is protein that increases or decreases transcription.

112. The method of claim 111, wherein the heterologous polypeptide is a transcriptional repressor domain.

113. The method of claim 111, wherein the heterologous polypeptide is a transcriptional activation domain.

114. The method of any one of claims 90, 91, 92, 96, or 97, wherein the heterologous polypeptide is a protein binding domain.

115. A transgenic, multicellular, non-human organism whose genome comprises a transgene comprising a nucleotide sequence encoding one or more of:

a) a CasP6 polypeptide, and/or

b) a CasP6 guide RNA.

116. The transgenic, multicellular, non-human organism of claim 115, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more amino acid sequence identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

117. The transgenic, multicellular, non-human organism of claim 115, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more amino acid sequence identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

118. The transgenic, multicellular, non-human organism of any one of claims 115-117, wherein the organism is a plant, a monocotyledon plant, a dicotyledon plant, an invertebrate animal, an insect, an arthropod, an arachnid, a parasite, a worm, a cnidarian, a vertebrate animal, a fish, a reptile, an amphibian, an ungulate, a bird, a pig, a horse, a sheep, a rodent, a mouse, a rat, or a non-human primate.

119. A system comprising:

a) a CasP6 polypeptide and a CasP6 single guide RNA;

b) a CasP6 polypeptide, a CasP6 guide RNA, and a DNA donor template;

c) a CasP6 fusion polypeptide and a CasP6 guide RNA;

d) a CasP6 fusion polypeptide, a CasP6 guide RNA, and a DNA donor template;

e) an mRNA encoding a CasP6 polypeptide, and a CasP6 single guide RNA;

f) an mRNA encoding a CasP6 polypeptide; a CasP6 guide RNA, and a DNA donor template;

g) an mRNA encoding a CasP6 fusion polypeptide, and a CasP6 guide RNA;

h) an mRNA encoding a CasP6 fusion polypeptide, a CasP6 guide RNA, and a DNA donor template;

i) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasP6 polypeptide; and ii) a nucleotide sequence encoding a CasP6 guide RNA;

j) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasP6 polypeptide; ii) a nucleotide sequence encoding a CasP6 guide RNA; and iii) a DNA donor template;

k) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasP6 fusion polypeptide; and ii) a nucleotide sequence encoding a CasP6 guide RNA; or

l) one or more recombinant expression vectors comprising: i) a nucleotide sequence encoding a CasP6 fusion polypeptide; ii) a nucleotide sequence encoding a CasP6 guide RNA; and a DNA donor template.

120. The CasP6 system of claim 119, wherein the CasP6 polypeptide comprises an amino acid sequence having 50% or more amino acid sequence identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

121. The CasP6 system of claim 119, wherein the CasP6 polypeptide comprises an amino acid sequence having 85% or more amino acid sequence identity to any of the amino acid sequences set forth in SEQ ID NO: 1, 2, or 3.

122. The composition of claim 13, wherein the DNA donor template has a length of from 8 nucleotides to 1000 nucleotides.

123. The composition of claim 13, wherein the DNA donor template has a length of from 25 nucleotides to 500 nucleotides.

124. The method of claim 87, wherein the DNA donor template has a length of from 8 nucleotides to 1000 nucleotides.

125. The method of claim 87, wherein the DNA donor template has a length of from 25 nucleotides to 500 nucleotides.

126. A nucleic acid detection system comprising: (i) a CasP6 polypeptide set forth in any one of claims 1, 2, or 9; (ii) one or more CasP6 guide RNAs each comprising a guide sequence that can hybridize to RNA molecules having corresponding target sequences; and (iii) an RNA-based masking construct.

127. The detection system of claim 126, wherein the RNA-based masking construct does not comprise the corresponding target sequence(s).

128. A method of detecting a target RNA molecule comprising:

(a) contacting target RNA molecule with: (i) a CasP6 polypeptide of claim 1, 2, or 9; or a CasP6 fusion polypeptide of any one of claims 24-26, 30-46, or 47; (ii) one or more CasP6 guide RNAs each comprising a guide sequence that can hybridize to the target RNA molecule; and (iii) an RNA-based masking construct, wherein the CasP6 comprises endonuclease enzyme activity, and wherein the RNA-based masking construct does not comprise the target sequence; and

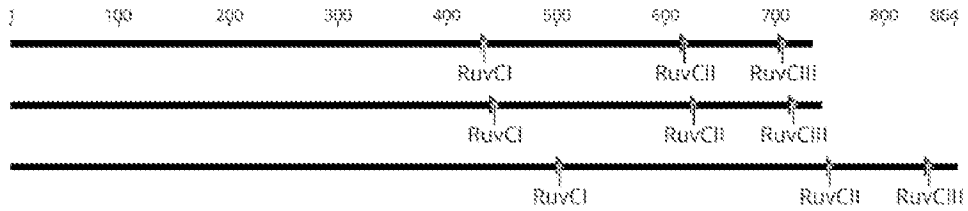
(b) assaying for modification of the RNA-based masking construct.

>CasP6A (SEQ ID NO: 1)
MITRLSLVNNCIHVHAGEISLKTTKAFTHGSLMFDQNTRLATFGKLPPVELPEGLTLTDGATMEFQLV
APDGTLRKQARAGVLKVSPKEPGIMRIAGMSKTPRAFSAADGWGIAIYKYRAYFSHPGVSTDAEIPE
WLKASIARQKAFWNRLAYLCREARRKCSPDSTTEVIAFVKETILPEIDAYNEALGRSKEKIKHPAKLK
TEMPGVDGLWNFVGILRGRIEKGRSVPDGLLEKVIAFAEQFKADYTPLNDFMNNYNDIAAKEAEALG
LRRFEIRPIVAGFKAVIDRRKTMKTTWSEGWPLLKYPDSPKADNWGLHYYFNMAGLDSAQFQSPKG
IPGLTLGPALKPSETGHPLLTGNAAKKNTLHEAEISISGTDREQWRFRFGVLCHRPLPAGSHLKEWKL
IFADGKLWLCLTVELQRPVPVHSPTAAGLDVGWRRTEEGIRFGTLYEPATKSFRELTIDMQRSPKDP
KVRVPFRIDVGPTRWEKRNITKLLPDWKPGDAIPSSFEIRPVVQSRRDYYKDTAKILLRKHLGEKLPA
WIDKAGRRGLLKVAEEFKEDTYVQGVIAEWKQQDEQIGNLVSMYLDRCTKRLEYGHAQVAHDVCR
HLQEKGVHRLSVEGSFLAKLSQRHDNEDAEALKRSQKYRQYAAVGKFIAVLKNIASKYGIVVDEISA
VNTTRICQHCNHLNPSSETEHFVCEECGRQVKQDHNASVNLSRFSCDPELAEMAVTAGKP

>CasP6B (SEQ ID NO: 2)
MLRVNPLPLRGGFCLPSSSPIASNYAQIFISCAAHRARSFKMIGRLTSEGSCIYVHCAEISLS
NTKAFTHGTLVVNLAARTVTFGKLPAVALPAGLVLAAGDTLKLQLVAPDGTLRKQTRIDVLK
ANEKQPGILRIAGSTTSERRFSAADGWGTAIYKYRAYFTHPGMNLEGELPQWLKGSIARQR
GLWNRLAWLCREARRKCSPVPTEEICAFVQETILPAIDAFNDSLGRSKQKMKHPAKLKTEA
PGIDGLWKFVGELRKRIEMDRAVPDGLLEKVIAFAEQFHPNYTPLNEFMNSFQEIAEREAGA
LVLEKDGAPVKLRPYEARPTIKSFKAVLDRRKTTKAPWSEGWPLIKFSDSPKAENWGLHFY
FNKAGVESSLLETDKGVPGLTFGPALLPAKTGHKKLVGDAAKKRRLREAEISIAGENGERW
DFKFGVLQHRPLPENSHVKEWKLVYQGHKLWLCLVVELQRPIPVISQQAAGLDIGWRRTEE
GIRFGTLYEPESKTFRELNIDLQKSPKDHSDRVPFRIDLGPTRWEKHNITALLPDWKPGDAIP
SAFEIRPAMNVRRSYYKDTAKILLRKHLSDRVPVWFDKAGSKGLFKLGEEFKEDATVQKILS
DWRQKEEEIGKLMAMYFDRTTLRVEYGHAQVAHDVCRYLMQKGTTRLIVETSFLARASQQ
HDNEDPDSLKNSQKYRQFAAVGKFVTILRNIAAKKGIVVDAHEAINTTRICQYCNHLNPSTEK
DQFSCEACGRQIKQDQNASVNLSRFGTDPDLAEMALLAGKA

>CasP6C (SEQ ID NO: 3)
MFSKLEFLGSSIQDGQETVSLGLHAAEVDNKLAQTSRFKAASVLIDAAAQVVRVREGSAPLPLPTEFT
PRKGQVIHLSLFDVEGKPRLQPRVGTLNAVKDATTKEDSLTEFRLPGQSSTPAQYSAEDGWGHRTL
KFRAYIKHTGVETDGAIPLWVERSVIRQRLARNYVAYLCREARRACSSVSVEQIIQFRDEVLKPELAA
FNAELDQKRADGIRLGVKGAKSFNKGKIPLPKQIDAESPEIASLLSFANKLFKLTKVNEELVNKGEAD
KQRPVPVALMTKIYAFRDQFPVNYDPINEFQSKALSTEFWKGITEKFGLSFYEVEFLSKRLRGDLERR
KTNGDTFDKGWPKIRYEDDSKFHDWTLYFRLGKPGVDARGLLSEQGIRTLRLGPPVSSQDSGHSE
MKQGSGKSLRALRQAEISLPFRVDNKRVRRTFTFAVLAHEEIPEGAHFVKWELNYKKVRKGSAHQS
PGGKRGSGYELFLTITLKVKNPTIGDIDQSAGLDIGWYKEPNGLRIGTLFLPAEHRYQKITLNLEESAP
DANTPFHLNVGASRWGRRYAQRAANELAVIAEEVDISAALFGIDSSKAGLPFEEKVSLSLKNLALAV
KGSEKGEEALPTIDIADAITRFRSLVRISHAFPFVDDFNGKEALQTRRDVLKDAAKAIIARQLGSDAP
AWLAKAGITGIRSIAEETNHPEALETISVWSDVDDLFGERAAFIAEKIADRLKKGYERLAHDICSKLA
GNAQCLGFEEDFLKKLGEAKSNYDHAALKASAKYRHFAALGSFVETVKRIAVKYAIRTIAVQAPGST
NTCNVCGHQNEVGVDRVYQCGDCGQTIDQDENAAVNLATNAFDVVQQETERNAVAA

FIG 1



RuvC Catalytic residues:

CasP6A	RuvC-I: AGLDVGWR (SEQ ID NO: 10)
	RuvC-II: LSVEGSFL (SEQ ID NO: 11)
	RuvC-III: QDHNASVNL (SEQ ID NO: 12)
CasP6B	RuvC-I: AGLDIGWR (SEQ ID NO: 13)
	RuvC-II: LIVETSFL (SEQ ID NO: 14)
	RuvC-III: QDQNASVNL (SEQ ID NO: 15)
CasP6C	RuvC-I: AGLDIGWY (SEQ ID NO: 16)
	RuvC-II: LGFEEDFL (SEQ ID NO: 17)
	RuvC-III: QDENAAVNL (SEQ ID NO: 18)

FIG 2

DR Sequences:

CasP6A- GTTGCAACGAGGTTGAAGTGCATCGAATAGATGCAAC (SEQ ID NO: 4)

CasP6B- CGTGGCAATGCCTCACAAGTGCACCGCATAGGTGCAAC (SEQ ID NO: 5)

CasP6C- CCTGAAGTAGAGCTGAGTGTCTCGGGTAGAGACAAC (SEQ ID NO: 6)

FIG 3A

CasP6A RNA
(SEQ ID NO: 7)

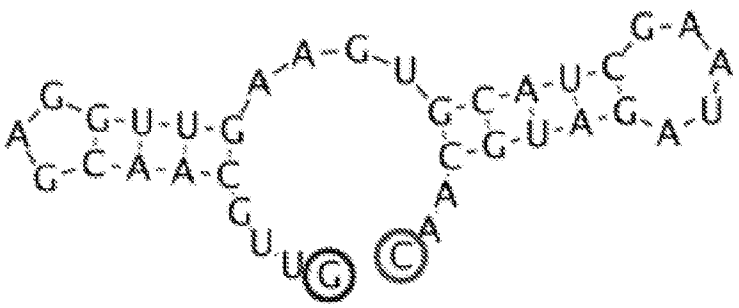


FIG. 3B

CasP6B RNA
(SEQ ID NO: 8)

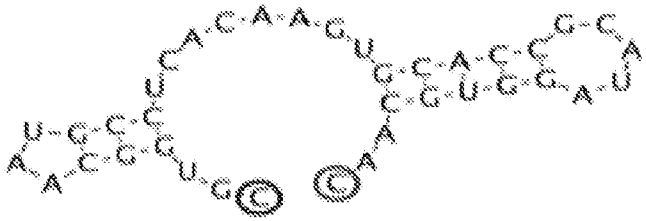


FIG. 3C

CasP6C RNA
(SEQ ID NO: 9)

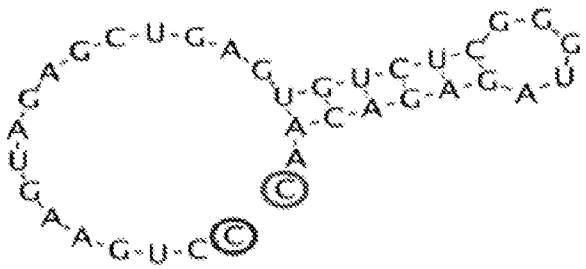


FIG. 3D

	1	2	3	4	5	6	7	8
Protein	+	+	+	+	+	+	+	+
Guide	-	+	-	+	-	+	-	+
Temp (C)	27	27	32	32	37	37	42	42

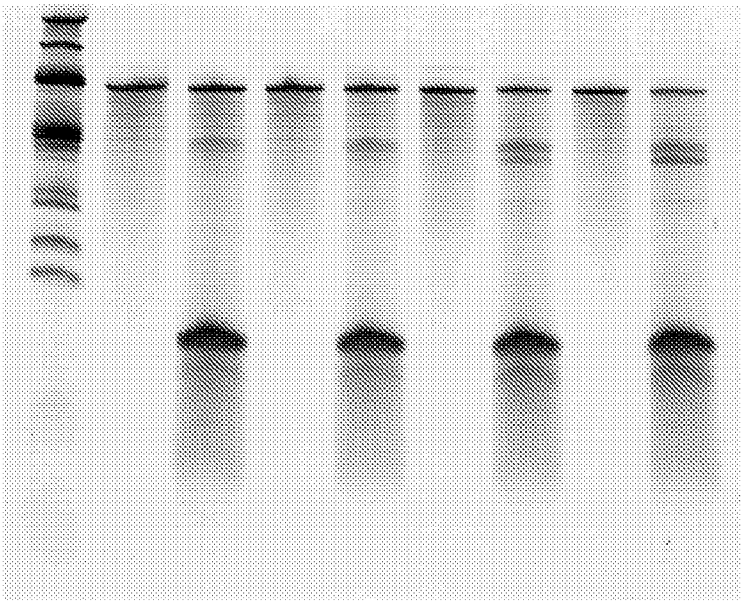


FIG. 4A

	1	2	3	4	5	6
Protein	+	+	+	+	+	+
Guide	-	+	+	+	+	+
Temp (C)	37	42	47	52	57	62

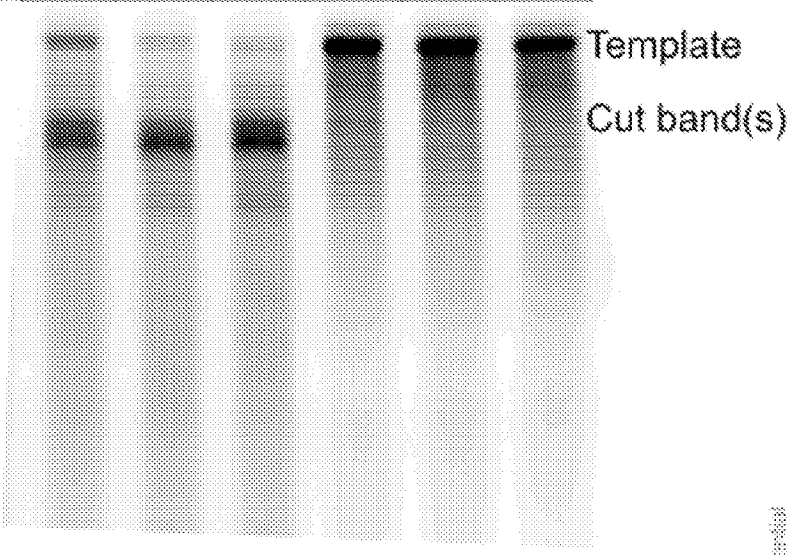


FIG. 4B

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/60442

A. CLASSIFICATION OF SUBJECT MATTER

IPC - C12N 9/22; C12N 15/81; C12N 15/86; C07K 19/00; C12N 15/90; C07K 14/195 (2020.01)

CPC - C12N 2310/20; C07K 2319/00; C12N 9/22; C12N 15/81; C12N 15/86; C12N 15/902; C12Y 2301/21004; C07K 14/195

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
See Search History document

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 ---- Y ---- A	US 2018/0119121 A1 (CARIBOU BIOSCIENCES INC.) 3 May 2018 (03.05.2018). Especially para [0046], [0052], [0084], [0085], [0086], [0091], [0093], [0095], [0101], [0104],, [0106-0107], [0170], [0178], claim 47.	1, 5-8, 14, 24, (30-31)/24, 48/24, 56/48, 70-73, 75/3, 115, 118/115 ---------- 3, 13/1, (16-19)/14, 49-51, 53, 54, 57, (60,63-68)/57, 119, 122, 123, (126,127)/(1,2) ---------- 2, 4, 9, 15, (16-19)/15, 25, 26, 58, 59, 60/(58,59), (61,62)/(57-59), (63-68)/(58,59), 116, 117, 118/(116,117), 120, 121, 126/9, 127/9
Y	WO 2018/130830 A1 (OXFORD UNIVERSITY INNOVATION LTD.) 19 July 2018 (19.07.2018). Especially pg 2 ln 13-18, p 44 ln 28, claims 1, 3-5	3, (16-19)/14, 119, (126,127)/(1,2)

☒ 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
"D" document cited by the applicant in the international application
"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	Date of mailing of the international search report
2 March 2020	30 MAR 2020
Name and mailing address of the ISA/US Mail Stop PCT, Attn: ISA/US, Commissioner for Patents P.O. Box 1450, Alexandria, Virginia 22313-1450 Facsimile No. 571-273-8300	Authorized officer Lee Young Telephone No. PCT Helpdesk: 571-272-4300

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/60442

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2018/0320202 A1 (THE REGENTS OF THE UNIVERSITY OF CALIFORNIA) 8 November 2018 (08.11.2018). Especially para [0004], [0128]	13/1, 122, 123
Y ---- A	WO 2018195555 A1 (THE LELAND STANFORD JUNIOR UNIVERSITY) 25 October 2018 (25.10.2018). Especially para [0016], [0094], [0099], [00120]	49-51, 53, 54, 57, (60,63-68)/57 ---------- 58, 59, 60/(58,59), (61,62)/(57-59), (63-68)/(58,59)
A	WO 2018/035250 A1 (THE BROAD INSTITUTE, INC.) 22 February 2018 (22.02.2018) Especially sheets 9-11 Fig 9 CRISPR protein 118577413.	2, 9, 25, 26, 58, 59, 60-68/(58,59) 116, 117, 118/(116,117) 120, 121, 126/9
A	US 2007/020621 A1 (BOUKHAROV et al.) 25 January 2007 (25.01.2007). Especially SEQ ID NO: 38063.	4, 15, (16-19)/15, (61,62)/57

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/60442

Box No. I 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 13*ter*.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 13*ter*.1(a)).

      ☐ on paper or in the form of an image file (Rule 13*ter*.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:

GenCore ver 6..4.1 SEQ ID NOs: 1, 2, 4, 5, 6 were searched.

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 19/60442

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.: 20-22, 52, 128
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:
-----Go to Extra Sheet for continuation------

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.:
1-9, 13 (in part), 14-19, 24-26, (30,31,48-51, 53, 54,56) (in part), 57-68, 70-73, 75 (in part), 115-123, 126-127 limited to Cas6P polypeptide is a ds DNA nuclease, has at least 50% amino acid identity to SEQ ID NO: 1, Cas6P guide RNA is encoded by residues 17-37 of SEQ ID NO: 4, is operably linked to a eukaryotic promoter, the first named heterologous fusion, nuclear localization signal (NLS).

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
Information on patent family members

International application No.
PCT/US 19/60442

Continuation of Box III: Observations where Unity of Invention is laking

This application contains the following inventions or groups of inventions which are not so linked as to form a single general inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I+: Claims 1-19, 23-51, 53-75, 115-123, 126-127, drawn to a composition or system comprising CasP6 polypeptide and a CasP6 guide RNA.

The CasP6 polypeptide and CasP6 guide RNA will be searched to the extent that the CasP6 polypeptide has double-stranded DNA nuclease activity (see instant application para [0042]) and the CasP6 polypeptide has an amino acid sequence having at least 50% sequence identity to SEQ ID NO: 1 (claim 2) and the CasP6 guide RNA encoded by residues 17-37 of SEQ ID NO: 4 (claim 4) with P6 guide RNA operably linked to a promoter functional in a eukaryotic cell (claim 18), and the first named heterologous fusion, nuclear localization signal (NLS)(claim 31). It is believed that claims 1-9, 13 (in part), 14-19, 24-26, (30,31,48-51, 53, 54,56) (in part), 57-68, 70-73, 75 (in part), 115-123, 126-127 read on this first named invention and thus these claims will be searched without fee to the extent that they encompass CasP6 with double stranded DNA nuclease activity, CasP6 has at least 50% sequence identity to SEQ ID NO: 1, and the CasP6 guide RNA encoded by residues 17-37 of SEQ ID NO: 5 with a promoter functional in a eukaryotic cell, and the heterologous fusion is NLS. Additional CasP6 polypeptides, CasP6 guide RNAs, and CasP6 guide RNA promoters, and heterologous fusions will be searched upon payment of additional fees. Applicant must specify the claims that encompass any additional CasP6 polypeptides, CasP6 guide RNAs, and CasP6 guide RNA promoters and heterologous fusions. Applicants must further indicate, if applicable, the claims which read on the first named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the "+" group(s) will result in only the first claimed invention to be searched/examined. An exemplary election would be: CasP6 polypeptide with at least 50% amino acid sequence identity to SEQ ID NO: 2, the CasP6 polypeptide is catalytically inactive (claim 11), and with P6 guide RNA operably linked to a promoter functional in a eukaryotic cell, and heterologous fusion is a nuclease (claims 1-9, 11, (12,13)(in part) 14-19, 24-26, 28, 29(in part), 34, (48-54,56)(in part), 57-68, 70-73, 75 (in part), 115-123, 126, 127).

Group II+: Claims 76-114, 124-125, drawn to a method of modifying a target molecule.

Group II+ will be searched upon payment of additional fee(s). The method may be searched, for example, to the extent that the CasP6 polypeptide cleaves a target, ds DNA (claims 77,78), and the cell is a eukaryotic cell (claim 82), the CasP6 guide RNA encoded by residues 17-37 of SEQ ID NO: 4 (claim 92) for an additional fee and election as such. It is believed that claims 77-83, 85-89, 90-92, 94-99, 101, 103-107, (124-125)(in part) read on this exemplary invention. Additional CasP6 polypeptide constructs, activities, cell type and CasP6 guide RNA sequences will be searched upon the payment of additional fees. Applicants must indicate, if applicable, which claims read on this named invention if different than what was indicated above for this group. Failure to clearly identify how any paid additional invention fees are to be applied to the '+' group(s) will result in only the first named invention to be searched/examined. An exemplary election would be the CasP6 polypeptide comprises a heterologous fusion with deamination activity (claim 105), the target activity is dsDNA (claim 78), the cell is a eukaryotic cell, the CasP6 guide RNA encoded by residues 17-37 of SEQ ID NO: 4 (claim 92) (claims 76, 78(in part), 80-83, (85-89)(in part), 90-92, 94-99, 101, 103-107, (124-125)(in part)).

The inventions listed as Groups I+ and II+ do not relate to a single general inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons:

Technical Features:

Group I+ has the special technical feature of a composition comprising CasP6 polypeptide and a CasP6 guide RNA, not required by Group II+.

Group II has the special technical feature of a method of modifying a target molecule, not required by Group I+.

No technical features are shared between the endonuclease polypeptide and polynucleotide sequences of Group I+ and, accordingly, this groups lacks unity a priori.

Additionally, even if Groups I+ inventions and Group II were considered to share the technical features of:

1. Groups I+ and Group II+ share the technical feature of a CasP6 polypeptide and a CasP6 guide RNA.

2. Group I+ inventions share claims 1, 14, 24, 70, 73, 115, 119.

3. Group II+ inventions share claims 76 and 90.

these shared technical features are previously disclosed by US 2018/0119121 A1 to Caribou Biosciences, Inc. (hereinafter "Caribou) [published 3 May 2018], in view of US 2018/0273962 A1 to Monsanto Technology LLC (hereinafter "Monsanto") [published 27 September 2018]

----continued on next sheet------

----continued from previous sheet------

As to claim 1, Caribou discloses a composition comprising:
a) a CasP6 polypeptide, or a nucleic acid molecule encoding the CasP6 polypeptide (claim 47); and
b) a CasP6 guide RNA, or one or more DNA molecules encoding the CasP6 guide RNA (claim 47; "47. A Type I CRISPR-Cas system protein complex, comprising: an artificial fusion protein comprising a FokI fused to the N-terminus of a Cse1 protein subunit, wherein the designed fusion protein further comprises a linker polypeptide between the FokI and the Cse1 protein subunit; a Cas6 protein subunit; a Cas5 protein subunit; a Cse2 protein subunit; a Cas7 protein subunit; and a CRISPR-derived RNA (crRNA) comprising a spacer sequence complementary to a target nucleic acid").

As to claim 14, Caribou discloses a CasP6 single guide RNA molecule, comprising:
a) a guide sequence that hybridizes to a target nucleic acid, and a duplex-forming segment (para [0010]; "The crRNA serves as a guide RNA for sequence specific binding of Cascade to double stranded (ds) DNA molecules through base pairing between the crRNA spacer and the complementary protospacer"); and
b) an activator sequence that can bind a CasP6 polypeptide (para [0010]; "there is a cleavage of the pre-crRNA transcript by the Cas6e subunit of Cascade, resulting in the mature 61 nt crRNA being retained by the CRISPR complex").

As to claim 24, Caribou discloses a CasP6 fusion polypeptide comprising: a CasP6 polypeptide fused to a heterologous polypeptide (claim 47)(para [0046]; "preferably the at least one functional moiety is fused or linked to at least the region of the N terminus and/or the region of the C terminus of at least one of a Cse1, Cse2, Cas7, Cas5, Cas6 or Cas6e subunit").

As to claim 70, Caribou discloses a eukaryotic cell (para [0104]; "In any of the methods of the invention as described above, the cell which is subjected to the method may be a prokaryote. Similarly, the cell may be a eukaryotic cell") comprising one or more of:
a) a CasP6 polypeptide, or a nucleic acid molecule encoding the CasP6 polypeptide (claim 47),
b) a CasP6 guide RNA, or a nucleic acid molecule encoding the CasP6 guide RNA (claim 47).

As to claim 73, Caribou discloses cell (para 0104]) comprising a CasP6 fusion polypeptide (claim 47; para [0046]).

As to claim 119, Caribou discloses a system comprising: a) a CasP6 polypeptide and a CasP6 single guide RNA (claim 47).

As to claim 115, Monsanto discloses a transgenic, multicellular, non-human organism whose genome comprises a transgene comprising a nucleotide sequence encoding one or more of:
a) a CasP6 polypeptide (claim 19; "19. The method of claim 18, wherein the site-specific genome modification enzyme is: ... b. an endonuclease selected from the group comprising ... Cas6"; claim 9; "9. A method of making a transgenic maize cell comprising at least one DNA of interest integrated in a B chromosome: i. selecting at least one target site having at least 90% sequence identity to a portion of a maize B chromosome sequence selected from the group consisting of SEQ ID NO:1-126 and 128-888; ii. selecting a site specific genome modification enzyme that specifically cleaves the target site; iii. introducing the site specific genome modification enzyme into the maize cell; iv. introducing the DNA of interest into the maize cell; v. integrating the DNA of interest into the target site; and vi. selecting a transgenic maize cell comprising the at least one of the DNA of interest integrated into the B chromosome").

As to claim 76, concerning a method of modifying a target nucleic acid, the method comprising contacting the target nucleic acid, Caribou discloses:
a) a CasP6 polypeptide (claim 47); and
b) a CasP6 guide RNA comprising a guide sequence that hybridizes to a target sequence of the target nucleic acid (claim 47),
wherein said contacting results in modification of the target nucleic acid by the CasP6 polypeptide (abstract; "Targeted ribonucleoproteins can be used as genetic engineering tools for precise cutting of nucleic acids in homologous recombination, non-homologous end joining, gene modification, gene integration, mutation repair or for their visualisation, transcriptional activation or repression").

As to claim 90, Caribou discloses a method of modulating transcription from a target DNA (abstract), modifying a target nucleic acid (abstract), or modifying a protein associated with a target nucleic acid (abstract), the method comprising contacting the target nucleic acid with:
a) a CasP6 fusion polypeptide comprising a CasP6 polypeptide fused to a heterologous polypeptide (claim 47); and
b) a CasP6 guide RNA comprising a guide sequence that hybridizes to a target sequence of the target nucleic acid (claim 47).

As the shared technical features were known in the art at the time of the invention, they cannot be considered shared special technical features that would otherwise unify the groups. The inventions lack unity with one another.

Therefore, Groups I+ and II+ lack unity of invention under PCT Rule 13 because they do not share a same or corresponding special technical feature.

Item 1(iii) (continued): Will establish the international search report on those parts of the international application which relate to the invention first mentioned in claims 1-9, 13 (in part), 14-19, 24-26, (30,31,48-54,56) (in part), 57-68, 70-73, 75 (in part), 115-123, 126-127 limited to Cas6P polypeptide is a ds DNA nuclease, has at least 50% amino acid identity to SEQ ID NO: 1, Cas6P guide RNA is encoded by residues 17-37 of SEQ ID NO: 4, is operably linked to a eukaryotic promoter, the first named heterologous fusion, nuclear localization signal (NLS).

Item 4 (cont.): Claims 20-22, 52, 128 are held unsearchable because they are not drafted according to the second and third sentences of PCT Rule 6.4(a).