



(51) International Patent Classification:

C12N 15/113 (2010.01) C12N 5/079 (2010.01)

A61K 48/00 (2006.01) C12N 9/22 (2006.01)

A61P 27/02 (2006.01) C12N 15/62 (2006.01)

C07K 14/47 (2006.01) C12N 15/63 (2006.01)

(21) International Application Number:

PCT/US2023/064004

(22) International Filing Date:

09 March 2023 (09.03.2023)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

63/319,088 11 March 2022 (11.03.2022) US

(71) Applicant: EPICRISPR BIOTECHNOLOGIES, INC.

[US/US]; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US).

(72) Inventors: COLLIN DE L'HORTET, Alexandra,

Sylvie; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US). TSOA, Rosemarie, Wenting; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US). NEUBAUER, Alexander, A.; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US). SALZMAN, Amber, Ruth; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US). HART, Daniel; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US). GAINOUS, Thomas; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US). CAROSSO, Giovanni; 7000

Shoreline Court, Suite 100, South San Francisco, California 94080 (US). YEO, Robin, W.; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US). DALEY, Timothy, Patrick; 7000 Shoreline Court, Suite 100, South San Francisco, California 94080 (US).

(74) Agent: BARKER, Adrian; Wilson Sonsini Goodrich & Rosati, 650 Page Mill Road, Palo Alto, California 94304 (US).

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

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

(54) Title: SYSTEMS AND METHODS FOR GENETIC MODULATION TO TREAT OCULAR DISEASES

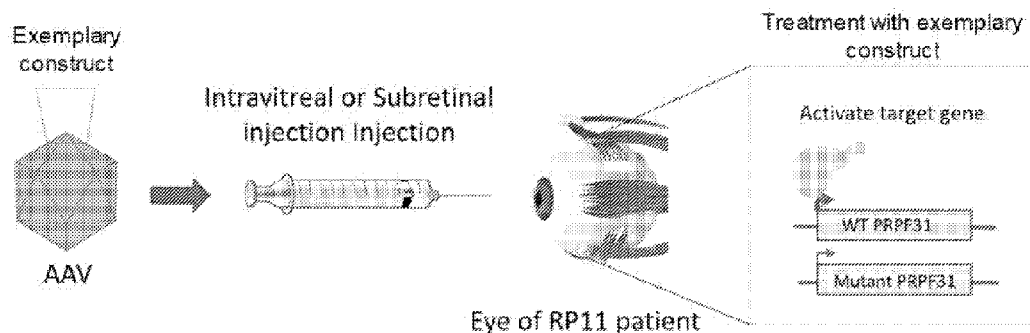


FIG. 2

(57) Abstract: Described herein are systems and methods for modulating gene expression. Also described herein are systems and methods for treating a disease or a condition by modulating gene expression.

Published:

— *with international search report (Art. 21(3))*

SYSTEMS AND METHODS FOR GENETIC MODULATION TO TREAT OCULAR DISEASES

CROSS-REFERENCE

[0001] This application claims the benefit of U.S. Provisional Application No. 63/319,088, filed March 11, 2022, which is incorporated herein by reference in its entirety.

BACKGROUND

[0002] Aberrant expression of one or more genes (e.g., endogenous genes) can lead to a disease or a condition in a subject. In some cases, aberrant expression of an enzyme regulator (e.g., an enzyme inhibitor, such as a protease inhibitor) in a cell in the subject can lead to irregular enzymatic activity within the cell, thereby effecting various diseases. The aberrant expression can be due to one or more hereditary genetic mutations in a gene encoding the enzyme regulator. For example, mutation of PRPF31 can lead to retinitis pigmentosa or autosomal congenital blindness in a subject.

SUMMARY

[0003] Modifying aberrant expression of a mutant allele (e.g., a disease-causing allele) in a cell may not be sufficient to treat or cure a disease that is manifested by the aberrant expression of the mutant allele. Thus, there remains a substantial need for systems, compositions, and methods to modify the aberrant expression of the mutant allele and introduce expression of a non-disease causing allele (e.g., a wild-type allele) in a cell.

[0004] In an aspect, the present disclosure provides a system comprising: a heterologous polypeptide comprising an actuator moiety, wherein the actuator moiety is for binding an endogenous target gene encoding PRPF31 in a cell, to increase expression level of the PRPF31 in the cell, wherein: (i) the actuator moiety substantially lacks DNA cleavage activity; and/or (ii) the actuator moiety is coupled to a transcriptional activator.

[0005] In another aspect, the present disclosure provides one or more polynucleotides encoding any of the system provided herein.

[0006] In another aspect, the present disclosure provides a method comprising administering any of the system or the one or more polynucleotides provided herein to a subject in need thereof.

[0007] In another aspect, the present disclosure provides a method comprising: increasing expression level of an endogenous target gene encoding PRPF31 in a cell, via binding of a heterologous polypeptide comprising an actuator moiety to bind the endogenous target gene, wherein: (a) the actuator moiety substantially lacks DNA cleavage activity; and/or (b) the actuator moiety is coupled to a transcriptional activator.

[0008] Additional aspects and advantages of the present disclosure will become readily apparent to those skilled in this art from the following detailed description, wherein only illustrative embodiments of the present disclosure are shown and described. As will be realized, the present disclosure is capable of other and different embodiments, and its several details are capable of modifications in various obvious respects, all without departing from the disclosure. Accordingly, the drawings and description are to be regarded as illustrative in nature, and not as restrictive.

INCORPORATION BY REFERENCE

[0009] All publications, patents, and patent applications mentioned in this specification are herein incorporated by reference to the same extent as if each individual publication, patent, or patent application was specifically and individually indicated to be incorporated by reference. To the extent publications and patents or patent applications incorporated by reference contradict the disclosure contained in the specification, the specification is intended to supersede and/or take precedence over any such contradictory material.

BRIEF DESCRIPTION OF THE DRAWINGS

[0010] The novel features of the invention are set forth with particularity in the appended claims. A better understanding of the features and advantages of the present invention will be obtained by reference to the following detailed description that sets forth illustrative embodiments, in which the principles of the invention are utilized, and the accompanying drawings (also “Figure” and “FIG.” herein), of which:

[0011] FIG. 1 illustrates an exemplary construct encoding the dCas and the actuator moiety (effector). Legend: Promoter: photoreceptor specific or ubiquitous promoter; dCas: small dead Cas molecule such as dCasMini or equivalent effector: effector to activate expression such as VPR or equivalent thereof; Pr: Promoter for gRNA such as H1 or U6 or equivalent thereof; gRNA: gRNA targeting the endogenous target gene comprising PRPF31.

[0012] FIG. 2 illustrates a schematic for treating retinitis pigmentosa 11 (RP11) with the system described herein. AAV can be engineered to deliver an exemplary construct via subretinal or intravitreal injection to a subject in need thereof, where the expression of the exemplary construct can activate expression of PRPF31 encoded by the heterologous CDS of the construct.

[0013] FIG. 3 illustrates exemplary genomic loci NM_015629 (ENST00000419967.5) and (ENST00000321030.9) transcripts that can be targeted by the gRNA of the system and the method described herein.

[0014] **FIGs. 4A-4F** schematically illustrate example vectors encoding the system of the present disclosure.

[0015] **FIG. 5A** schematically illustrates an example experimental procedure to screen for target endogenous polynucleotide sequences to modulate expression of PRPF31, and **FIG. 5B** shows change in endogenous PRPF31 by different guide nucleic acid molecules.

[0016] **FIG. 6A** shows a fluorescent image of normal retinal pigment epithelium (RPE) cells, and **FIG. 6B** shows a fluorescent image of diseased RPE cells.

DETAILED DESCRIPTION

[0017] While various embodiments of the invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. Numerous variations, changes, and substitutions may occur to those skilled in the art without departing from the invention. It should be understood that various alternatives to the embodiments of the invention described herein may be employed.

[0018] Whenever the term “at least,” “greater than,” or “greater than or equal to” precedes the first numerical value in a series of two or more numerical values, the term “at least,” “greater than” or “greater than or equal to” applies to each of the numerical values in that series of numerical values. For example, greater than or equal to 1, 2, or 3 is equivalent to greater than or equal to 1, greater than or equal to 2, or greater than or equal to 3.

[0019] Whenever the term “no more than,” “less than,” or “less than or equal to” precedes the first numerical value in a series of two or more numerical values, the term “no more than,” “less than,” or “less than or equal to” applies to each of the numerical values in that series of numerical values. For example, less than or equal to 3, 2, or 1 is equivalent to less than or equal to 3, less than or equal to 2, or less than or equal to 1.

[0020] The term “about” or “approximately” generally mean within an acceptable error range for the particular value as determined by one of ordinary skill in the art, which will depend in part on how the value is measured or determined, i.e., the limitations of the measurement system. For example, “about” can mean within 1 or more than 1 standard deviation, per the practice in the art. Alternatively, “about” can mean a range of up to 20%, up to 10%, up to 5%, or up to 1% of a given value. Alternatively, particularly with respect to biological systems or processes, the term can mean within an order of magnitude, preferably within 5-fold, and more preferably within 2-fold, of a value. Where particular values are described in the application and claims, unless otherwise stated, the term “about” meaning within an acceptable error range for the particular value should be assumed.

[0021] The use of the alternative (e.g., “or”) should be understood to mean either one, both, or

any combination thereof of the alternatives. The term “and/or” should be understood to mean either one, or both of the alternatives.

[0022] The term “cell” generally refers to a biological cell. A cell can be the basic structural, functional and/or biological unit of a living organism. A cell can originate from any organism having one or more cells. Some non-limiting examples 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, tomatoes, rice, cassava, sugarcane, pumpkin, hay, potatoes, cotton, cannabis, tobacco, flowering plants, conifers, gymnosperms, ferns, clubmosses, hornworts, liverworts, mosses), 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., a pig, a cow, a goat, a sheep, a rodent, a rat, a mouse, a non-human primate, a human, etc.), and etcetera. Sometimes a cell is not originating from a natural organism (e.g. a cell can be a synthetically made, sometimes termed an artificial cell).

[0023] The term “nucleotide,” as used herein, generally refers to a base-sugar-phosphate combination. A nucleotide can comprise a synthetic nucleotide. A nucleotide can comprise a synthetic nucleotide analog. Nucleotides can be monomeric units of a nucleic acid sequence (e.g. deoxyribonucleic acid (DNA) and ribonucleic acid (RNA)). The term nucleotide can include ribonucleoside triphosphates adenosine triphosphate (ATP), uridine triphosphate (UTP), cytosine triphosphate (CTP), guanosine triphosphate (GTP) and deoxyribonucleoside triphosphates such as dATP, dCTP, dITP, dUTP, dGTP, dTTP, or derivatives thereof. Such derivatives can include, for example, $[\alpha S]$ dATP, 7-deaza-dGTP and 7-deaza-dATP, and nucleotide derivatives that confer nuclease resistance on the nucleic acid molecule containing them. The term nucleotide as used herein can refer to dideoxyribonucleoside triphosphates (ddNTPs) and their derivatives.

Illustrative examples of dideoxyribonucleoside triphosphates can include, but are not limited to, ddATP, ddCTP, ddGTP, ddITP, and ddTTP. A nucleotide may be unlabeled or detectably labeled by well-known techniques. Labeling can also be carried out with quantum dots. Detectable labels can include, for example, radioactive isotopes, fluorescent labels, chemiluminescent labels, bioluminescent labels and enzyme labels. Fluorescent labels of nucleotides may include but are not limited fluorescein, 5-carboxyfluorescein (FAM), 2'7'-dimethoxy-4'5'-dichloro-6-carboxyfluorescein (JOE), rhodamine, 6-carboxyrhodamine (R6G), N,N,N',N'-tetramethyl-6-

carboxyrhodamine (TAMRA), 6-carboxy-X-rhodamine (ROX), 4-(4'-dimethylaminophenylazo) benzoic acid (DABCYL), Cascade Blue, Oregon Green, Texas Red, Cyanine and 5-(2'-aminoethyl)aminonaphthalene-1-sulfonic acid (EDANS). Specific examples of fluorescently labeled nucleotides can include [R6G]dUTP, [TAMRA]dUTP, [R110]dCTP, [R6G] dCTP, [TAMRA] dCTP, [JOE] ddATP, [R6G] ddATP, [FAM] ddCTP, [R110]ddCTP, [TAMRA]ddGTP, [ROX]ddTTP, [dR6G]ddATP, [dR110]ddCTP, [dTAMRA]ddGTP, and [dROX]ddTTP available from Perkin Elmer, Foster City, Calif. FluoroLink DeoxyNucleotides, FluoroLink Cy3-dCTP, FluoroLink Cy5-dCTP, FluoroLink Fluor X-dCTP, FluoroLink Cy3-dUTP, and FluoroLink Cy5-dUTP available from Amersham, Arlington Heights, Ill.; Fluorescein-15-dATP, Fluorescein-12-dUTP, Tetramethyl-rhodamine-6-dUTP, IR770-9-dATP, Fluorescein-12-ddUTP, Fluorescein-12-UTP, and Fluorescein-15-2'-dATP available from Boehringer Mannheim, Indianapolis, Ind.; and Chromosome Labeled Nucleotides, BODIPY-FL-14-UTP, BODIPY-FL-4-UTP, BODIPY-TMR-14-UTP, BODIPY-TMR-14-dUTP, BODIPY-TR-14-UTP, BODIPY-TR-14-dUTP, Cascade Blue-7-UTP, Cascade Blue-7-dUTP, fluorescein-12-UTP, fluorescein-12-dUTP, Oregon Green 488-5-dUTP, Rhodamine Green-5-UTP, Rhodamine Green-5-dUTP, tetramethylrhodamine-6-UTP, tetramethylrhodamine-6-dUTP, Texas Red-5-UTP, Texas Red-5-dUTP, and Texas Red-12-dUTP available from Molecular Probes, Eugene, Oreg. Nucleotides can also be labeled or marked by chemical modification. A chemically-modified single nucleotide can be biotin-dNTP. Some non-limiting examples of biotinylated dNTPs can include, biotin-dATP (e.g., bio-N6-ddATP, biotin-14-dATP), biotin-dCTP (e.g., biotin-11-dCTP, biotin-14-dCTP), and biotin-dUTP (e.g. biotin-11-dUTP, biotin-16-dUTP, biotin-20-dUTP).

[0024] The term “polynucleotide,” “oligonucleotide,” or “nucleic acid,” as used interchangeably herein, generally refers to a polymeric form of nucleotides of any length, either deoxyribonucleotides or ribonucleotides, or analogs thereof, either in single-, double-, or multi-stranded form. A polynucleotide can be exogenous or endogenous to a cell. A polynucleotide can exist in a cell-free environment. A polynucleotide can be a gene or fragment thereof. A polynucleotide can be DNA. A polynucleotide can be RNA. A polynucleotide can have any three dimensional structure, and can perform any function, known or unknown. A polynucleotide can comprise one or more analogs (e.g. altered backbone, sugar, or nucleobase). If present, modifications to the nucleotide structure can be imparted before or after assembly of the polymer. Some non-limiting examples of analogs include: 5-bromouracil, peptide nucleic acid, xeno nucleic acid, morpholinos, locked nucleic acids, glycol nucleic acids, threose nucleic acids, dideoxynucleotides, cordycepin, 7-deaza-GTP, fluorophores (e.g. rhodamine or fluorescein linked

to the sugar), thiol containing nucleotides, biotin linked nucleotides, fluorescent base analogs, CpG islands, methyl-7-guanosine, methylated nucleotides, inosine, thiouridine, pseudouridine, dihydrouridine, queuosine, and wyosine. Non-limiting examples of polynucleotides include coding or non-coding regions of a gene or gene fragment, loci (locus) defined from linkage analysis, exons, introns, messenger RNA (mRNA), transfer RNA (tRNA), ribosomal RNA (rRNA), short interfering RNA (siRNA), short-hairpin RNA (shRNA), micro-RNA (miRNA), ribozymes, cDNA, recombinant polynucleotides, branched polynucleotides, plasmids, vectors, isolated DNA of any sequence, isolated RNA of any sequence, cell-free polynucleotides including cell-free DNA (cfDNA) and cell-free RNA (cfRNA), nucleic acid probes, and primers. The sequence of nucleotides can be interrupted by non-nucleotide components.

[0025] The term “sequence identity” generally refers to an exact nucleotide-to-nucleotide or amino acid-to-amino acid correspondence of two polynucleotides or polypeptide sequences, respectively. Typically, techniques for determining sequence identity include determining the nucleotide sequence of a polynucleotide and/or determining the amino acid sequence encoded thereby, and comparing these sequences to a second nucleotide or amino acid sequence. Two or more sequences (polynucleotide or amino acid) can be compared by determining their “percent identity.” The percent identity of two sequences, whether nucleic acid or amino acid sequences, is the number of exact matches between two aligned sequences divided by the length of the longer sequence and multiplied by 100. Percent identity may also be determined, for example, by comparing sequence information using the advanced BLAST computer program, including version 2.2.9, available from the National Institutes of Health. The BLAST program is based on the alignment method of Karlin and Altschul, *Proc. Natl. Acad. Sci. USA*, 87:2264-2268 (1990) and as discussed in Altschul, et al., *J. Mol. Biol.*, 215:403-410 (1990); Karlin And Altschul, *Proc. Natl. Acad. Sci. USA*, 90:5873-5877 (1993); and Altschul et al., *Nucleic Acids Res.*, 25:3389-3402 (1997). The program may be used to determine percent identity over the entire length of the proteins being compared. Default parameters are provided to optimize searches with short query sequences in, for example, with the blastp program. The program also allows use of an SEG filter to mask-off segments of the query sequences as determined by the SEG program of Wootton and Federhen, *Computers and Chemistry* 17:149-163 (1993). Ranges of desired degrees of sequence identity are approximately 50% to 100% and integer values therebetween. In general, this disclosure encompasses sequences with at least 50%, at least 55%, at least 60%, at least 65%, at least 70%, at least 75%, at least 80%, at least 85%, at least 90%, at least 95%, or at least 98% sequence identity with any sequence provided herein.

[0026] The term “gene” generally refers to a nucleic acid (e.g., DNA such as genomic DNA and

cDNA) and its corresponding nucleotide sequence that is involved in encoding an RNA transcript. The term as used herein with reference to genomic DNA includes intervening, non-coding regions as well as regulatory regions and can include 5' and 3' ends. In some uses, the term encompasses the transcribed sequences, including 5' and 3' untranslated regions (5'-UTR and 3'-UTR), exons and introns. In some genes, the transcribed region will contain "open reading frames" that encode polypeptides. In some uses of the term, a "gene" comprises only the coding sequences (e.g., an "open reading frame" or "coding region") necessary for encoding a polypeptide. In some cases, genes do not encode a polypeptide, for example, ribosomal RNA genes (rRNA) and transfer RNA (tRNA) genes. In some cases, the term "gene" includes not only the transcribed sequences, but in addition, also includes non-transcribed regions including upstream and downstream regulatory regions, enhancers and promoters. For example, a gene can refer to a portion of the gene that is near or adjacent to a transcription start site (TSS) of the gene. The gene (e.g., that is targeted as disclosed herein) can be at least or up to about 2,000 nucleobases, at least or up to about 1,800 nucleobases, at least or up to about 1,600 nucleobases, at least or up to about 1,500 nucleobases, at least or up to about 1,400 nucleobases, at least or up to about 1,200 nucleobases, at least or up to about 1,000 nucleobases, at least or up to about 900 nucleobases, at least or up to about 800 nucleobases, at least or up to about 700 nucleobases, at least or up to about 600 nucleobases, at least or up to about 500 nucleobases, at least or up to about 400 nucleobases, at least or up to about 300 nucleobases, at least or up to about 200 nucleobases, at least or up to about 100 nucleobases, or at least or up to about 50 nucleobases away from the TSS of the gene.

[0027] A gene can refer to an "endogenous gene" or a native gene in its natural location in the genome of an organism. A gene can refer to an "exogenous gene" or a non-native gene. A non-native gene can refer to a gene not normally found in the host organism but which is introduced into the host organism by gene transfer. A non-native gene can also refer to a gene not in its natural location in the genome of an organism. A non-native gene can also refer to a naturally occurring nucleic acid or polypeptide sequence that comprises mutations, insertions and/or deletions (e.g., non-native sequence).

[0028] The term "expression" generally refers to one or more processes by which a polynucleotide is transcribed from a DNA template (such as into an mRNA or other RNA transcript) and/or the process by which a transcribed mRNA is subsequently translated into peptides, polypeptides, or proteins. Transcripts and encoded polypeptides can be collectively referred to as "gene product." If the polynucleotide is derived from genomic DNA, expression can include splicing of the mRNA in a eukaryotic cell. "Up-regulated," with reference to

expression, generally refers to an increased expression level of a polynucleotide (e.g., RNA such as mRNA) and/or polypeptide sequence relative to its expression level in a wild-type state while “down-regulated” generally refers to a decreased expression level of a polynucleotide (e.g., RNA such as mRNA) and/or polypeptide sequence relative to its expression in a wild-type state.

Expression of a transfected gene can occur transiently or stably in a cell. During “transient expression” the transfected gene is not transferred to the daughter cell during cell division. Since its expression is restricted to the transfected cell, expression of the gene is lost over time. In contrast, stable expression of a transfected gene can occur when the gene is co-transfected with another gene that confers a selection advantage to the transfected cell. Such a selection advantage may be a resistance towards a certain toxin that is presented to the cell.

[0029] The term “expression profile” generally refers to quantitative (e.g., abundance) and qualitative expression of one or more genes in a sample (e.g., a cell). The one or more genes can be expressed and ascertained in the form of a nucleic acid molecule (e.g., an mRNA or other RNA transcript). Alternatively or in addition to, the one or more genes can be expressed and ascertained in the form of a polypeptide (e.g., a protein measured via Western blot). An expression profile of a gene may be defined as a shape of an expression level of the gene over a time period (e.g., at least or up to about 1 hour, at least or up to about 2 hours, at least or up to about 3 hours, at least or up to about 4 hours, at least or up to about 5 hours, at least or up to about 6 hours, at least or up to about 7 hours, at least or up to about 8 hours, at least or up to about 9 hours, at least or up to about 10 hours, at least or up to about 11 hours, at least or up to about 12 hours, at least or up to about 16 hours, at least or up to about 18 hours, at least or up to about 24 hours, at least or up to about 36 hours, at least or up to about 48 hours, at least up to about 3 days, at least up to about 4 days, at least up to about 5 days, at least up to about 6 days, at least up to about 7 days, at least up to about 8 days, at least up to about 9 days, at least up to about 10 days, at least up to about 11 days, at least up to about 12 days, at least up to about 13 days, at least up to about 14 days, etc.). Alternatively, an expression profile of a gene may be defined as an expression level of the gene at a time point of interest (e.g., the expression level of the gene measured at least or up to about 1 hour, at least or up to about 2 hours, at least or up to about 3 hours, at least or up to about 4 hours, at least or up to about 5 hours, at least or up to about 6 hours, at least or up to about 7 hours, at least or up to about 8 hours, at least or up to about 9 hours, at least or up to about 10 hours, at least or up to about 11 hours, at least or up to about 12 hours, at least or up to about 16 hours, at least or up to about 18 hours, at least or up to about 24 hours, at least or up to about 36 hours, at least or up to about 48 hours, at least up to about 3 days, at least up to about 4 days, at least up to about 5 days, at least up to about 6 days, at

least up to about 7 days, at least up to about 8 days, at least up to about 9 days, at least up to about 10 days, at least up to about 11 days, at least up to about 12 days, at least up to about 13 days, or at least up to about 14 days after treating a cell to induce such expression level.)

[0030] The term “peptide,” “polypeptide,” or “protein,” as used interchangeably herein, generally refers to a polymer of at least two amino acid residues joined by peptide bond(s). This term does not connote a specific length of polymer, nor is it intended to imply or distinguish whether the peptide is produced using recombinant techniques, chemical or enzymatic synthesis, or is naturally occurring. The terms apply to naturally occurring amino acid polymers as well as amino acid polymers comprising at least one modified amino acid. In some cases, the polymer can be interrupted by non-amino acids. The terms include amino acid chains of any length, including full length proteins, and proteins with or without secondary and/or tertiary structure (e.g., domains). The terms also encompass an amino acid polymer that has been modified, for example, by disulfide bond formation, glycosylation, lipidation, acetylation, phosphorylation, oxidation, and any other manipulation such as conjugation with a labeling component. The terms “amino acid” and “amino acids,” as used herein, generally refer to natural and non-natural amino acids, including, but not limited to, modified amino acids and amino acid analogues. Modified amino acids can include natural amino acids and non-natural amino acids, which have been chemically modified to include a group or a chemical moiety not naturally present on the amino acid. Amino acid analogues can refer to amino acid derivatives. The term “amino acid” includes both D-amino acids and L-amino acids.

[0031] The term “derivative,” “variant,” or “fragment,” as used herein with reference to a polypeptide, generally refers to a polypeptide related to a wild type polypeptide, for example either by amino acid sequence, structure (e.g., secondary and/or tertiary), activity (e.g., enzymatic activity) and/or function. Derivatives, variants and fragments of a polypeptide can comprise one or more amino acid variations (e.g., mutations, insertions, and deletions), truncations, modifications, or combinations thereof compared to a wild type polypeptide.

[0032] The term “engineered,” “chimeric,” or “recombinant,” as used herein with respect to a polypeptide molecule (e.g., a protein), generally refers to a polypeptide molecule having a heterologous amino acid sequence or an altered amino acid sequence as a result of the application of genetic engineering techniques to nucleic acids which encode the polypeptide molecule, as well as cells or organisms which express the polypeptide molecule. The term “engineered” or “recombinant,” as used herein with respect to a polynucleotide molecule (e.g., a DNA or RNA molecule), generally refers to a polynucleotide molecule having a heterologous nucleic acid sequence or an altered nucleic acid sequence as a result of the application of genetic engineering

techniques. Genetic engineering techniques include, but are not limited to, PCR and DNA cloning technologies; transfection, transformation and other gene transfer technologies; homologous recombination; site-directed mutagenesis; and gene fusion. In some cases, an engineered or recombinant polynucleotide (e.g., a genomic DNA sequence) can be modified or altered by a gene editing moiety.

[0033] The terms “engineered” and “modified” are used interchangeably herein. The terms “engineering” and “modifying” are used interchangeably herein. The terms “engineered cell” or “modified cell” are used interchangeably herein. The terms “engineered characteristic” and “modified characteristic” are used interchangeably herein.

[0034] The term “enhanced expression,” “increased expression,” or “upregulated expression” generally refers to production of a moiety of interest (e.g., a polynucleotide or a polypeptide) to a level that is above a normal level of expression of the moiety of interest in a host strain (e.g., a host cell). The normal level of expression can be substantially zero (or null) or higher than zero. The moiety of interest can comprise an endogenous gene or polypeptide construct of the host strain. The moiety of interest can comprise a heterologous gene or polypeptide construct that is introduced to or into the host strain. For example, a heterologous gene encoding a polypeptide of interest can be knocked-in (KI) to a genome of the host strain for enhanced expression of the polypeptide of interest in the host strain.

[0035] The term “enhanced activity,” “increased activity,” or “upregulated activity” generally refers to activity of a moiety of interest (e.g., a polynucleotide or a polypeptide) that is modified to a level that is above a normal level of activity of the moiety of interest in a host strain (e.g., a host cell). The normal level of activity can be substantially zero (or null) or higher than zero. The moiety of interest can comprise a polypeptide construct of the host strain. The moiety of interest can comprise a heterologous polypeptide construct that is introduced to or into the host strain. For example, a heterologous gene encoding a polypeptide of interest can be knocked-in (KI) to a genome of the host strain for enhanced activity of the polypeptide of interest in the host strain.

[0036] The term “reduced expression,” “decreased expression,” or “downregulated expression” generally refers to a production of a moiety of interest (e.g., a polynucleotide or a polypeptide) to a level that is below a normal level of expression of the moiety of interest in a host strain (e.g., a host cell). The normal level of expression is higher than zero. The moiety of interest can comprise an endogenous gene or polypeptide construct of the host strain. In some cases, the moiety of interest can be knocked-out or knocked-down in the host strain. In some examples, reduced expression of the moiety of interest can include a complete inhibition of such expression in the host strain.

[0037] The term “reduced activity,” “decreased activity,” or “downregulated activity” generally refers to activity of a moiety of interest (e.g., a polynucleotide or a polypeptide) that is modified to a level that is below a normal level of activity of the moiety of interest in a host strain (e.g., a host cell). The normal level of activity is higher than zero. The moiety of interest can comprise an endogenous gene or polypeptide construct of the host strain. In some cases, the moiety of interest can be knocked-out or knocked-down in the host strain. In some examples, reduced activity of the moiety of interest can include a complete inhibition of such activity in the host strain.

[0038] The term “subject,” “individual,” or “patient,” as used interchangeably herein, generally refers to a vertebrate, preferably a mammal such as a human. Mammals include, but are not limited to, murines, simians, humans, farm animals, sport animals, and pets. Tissues, cells and their progeny of a biological entity obtained in vivo or cultured in vitro are also encompassed.

[0039] The term “treatment” or “treating” generally refers to an approach for obtaining beneficial or desired results including but not limited to a therapeutic benefit and/or a prophylactic benefit. For example, a treatment can comprise administering a system or cell population disclosed herein. By therapeutic benefit is meant any therapeutically relevant improvement in or effect on one or more diseases, conditions, or symptoms under treatment. For prophylactic benefit, a composition can be administered to a subject at risk of developing a particular disease, condition, or symptom, or to a subject reporting one or more of the physiological symptoms of a disease, even though the disease, condition, or symptom may not have yet been manifested.

[0040] The term “effective amount” or “therapeutically effective amount” generally refers to the quantity of a composition, for example a composition comprising heterologous polypeptides, heterologous polynucleotides, and/or modified cells (e.g., modified stem cells), that is sufficient to result in a desired activity upon administration to a subject in need thereof. Within the context of the present disclosure, the term “therapeutically effective” generally refers to that quantity of a composition that is sufficient to delay the manifestation, arrest the progression, relieve or alleviate at least one symptom of a disorder treated by the methods of the present disclosure.

Overview

[0041] Gene expression underpins various physiological and pathological effects in cells and tissues, contributing to many diseases and conditions, thus agents that modulate expression of specific genes in a desirable way could have therapeutic benefit.

[0042] Developing agents that elicit robust, persistent, and/or reversible changes in gene expression has proven challenging, however, as many candidate therapeutics achieve only modest or short lived effects, or conversely result in off-target effects. Additionally, many current

approaches to gene editing and genome engineering can result in off-target effects that can be associated with undesirable toxicity profiles, and in some cases, undesirable effects can be permanent. There is thus a need for novel strategies to regulate gene expression that allow robust, persistent, and/or reversible modulation of target gene expression and activity, for example, expression of genes that impact human disease.

[0043] For instance, described herein are systems and methods for increasing expression of a target gene in a cell. In some embodiments, the target gene is encoded from a heterologous polynucleotide described herein. In some embodiments, the target gene encoded by the heterologous polynucleotide is a non-disease causing allele.

Systems, compositions, and methods thereof

[0044] In an aspect, the present disclosure provides a system comprising: a heterologous polypeptide comprising an actuator moiety, wherein the actuator moiety is for binding an endogenous target gene encoding PRPF31 in a cell, to increase expression level of the PRPF31 in the cell, wherein: the actuator moiety substantially lacks DNA cleavage activity; and/or the actuator moiety is coupled to a transcriptional activator. In some embodiments, described herein is one or more polynucleotides encoding a system described herein. Also described herein, in some aspects, is a method comprising administering the system described herein to a subject in need thereof. In some aspects, described herein is a method comprising: increasing expression level of an endogenous target gene encoding PRPF31 in a cell, via binding of a heterologous polypeptide comprising an actuator moiety to bind the endogenous target gene, wherein: the actuator moiety substantially lacks DNA cleavage activity; and/or the actuator moiety is coupled to a transcriptional activator.

[0045] In some embodiments, the systems, compositions, or methods described herein increase expression of PRPF31 in a cell. In some embodiments, the cell is an eye cell. In some embodiments, the cell is a retinal cell. In some embodiments, the cell is a retinal pigment epithelium (RPE) cell. In some embodiments, the PRPF31 expression being increased by the systems, compositions, or methods described herein is encoded from a non-disease allele of PRPF31. In some embodiments, the PRPF31 expression being increased by the systems, compositions, or methods described herein is encoded from a non-disease allele of PRPF31, where the non-disease allele of PRPF31 is encoded by a heterologous polynucleotide described herein. In some embodiments, the systems, compositions, or methods described herein increase expression of PRPF31 in a cell by contacting the cell with an actuator moiety described herein. In some embodiments, the actuator moiety is not capable of binding an additional target endogenous gene encoding a mutant allele of the PRPF31.

[0046] In some embodiments, described herein is a system for modulating a gene expression of an endogenous target gene described herein. In some embodiments, the system comprises a heterologous polypeptide comprising an actuator moiety, wherein the actuator moiety is for binding an endogenous target gene encoding a target protein in a cell, to decrease expression level of the target protein, and wherein the actuator moiety substantially lacks DNA cleavage activity; and a heterologous polynucleotide encoding a non-disease causing variant of the endogenous target gene. In some embodiments, the endogenous target gene is a non-disease causing variant. In some embodiments, the non-disease causing variant is a wild type variant. In some embodiments, the endogenous target gene is a disease causing variant. In some embodiments, the system comprises the heterologous polynucleotide not integrated into the endogenous target gene. In some embodiments, the system comprises the heterologous polypeptide that is under the control of a tissue-specific promoter. For example, the tissue-specific promoter can be a rod cell specific promoter, a cone cell specific promoter, a retina cell specific promoter, a photoreceptor specific promoter, or a combination thereof. In some embodiments, the tissue-specific promoter is a photoreceptor specific promoter. In some embodiments, the tissue-specific promoter is a PRPF31 promoter. Non-limiting example of tissue-specific promoter can include MOPS promoter, GRK1 promoter, IRBP promoter, PR2.1 promoter, IRBP/GNAT2 promoter, VMD2 promoter, VEcad/VEcadherin promoter, or a combination thereof.

[0047] In some embodiments, the system comprises the heterologous polypeptide that is under the control of a constitutive promoter. Non-limiting example of constitutive promoter can include CMV promoter, EF1a promoter, CAG promoter, PGK promoter, TRE promoter, U6 promoter, or UAS promoter. For example, the constitutive promoter can be a Pol III promoter (e.g., 7SK, U6, H1, etc.). In another example, the constitutive promoter can be a Pol II promoter (e.g., CMV, RSV, etc.). In some embodiments, the actuator moiety comprises a nuclease such as an endonuclease (e.g., a heterologous endonuclease). In some embodiments, the nuclease can be a deactivated nuclease such as a deactivated endonuclease, where the deactivated endonuclease does not cleave nucleic acid.

[0048] In some embodiments, the system comprises a guide nucleic acid. In some embodiments, a guide nucleic acid capable of forming a complex with the actuator moiety, wherein the complex binds the endogenous target gene. In some embodiments, the guide nucleic acid comprises a plurality of different guide nucleic acids capable of targeting different regions of the endogenous target gene.

[0049] In some embodiments, the system comprises an actuator moiety or a heterologous polynucleotide encoding the actuator moiety. In some embodiments, the actuator moiety is coupled to a transcriptional repressor. In some embodiments, the actuator moiety is fused to the transcriptional repressor.

[0050] In some embodiments, the system modulates a gene expression of an endogenous target gene in a cell. In some embodiments, the cell is an eye cell such as a rod cell, a cone cell, or a retina cell. In some embodiments, the cell is a photoreceptor cell, a bipolar cell, a retinal ganglion cell, a horizontal cell, or an amacrine cells. In some embodiments, the cell is a cell of pigmented layer. In some embodiments, the cell is a cell of layer of rods and cones. In some embodiments, the cell is a cell of membrana limitans externa. In some embodiments, the cell is a cell of outer nuclear layer. In some embodiments, the cell is a cell of outer plexiform layer. In some embodiments, the cell is a cell of inner nuclear layer. In some embodiments, the cell is a cell of inner plexiform layer. In some embodiments, the cell is a cell of ganglionic layer. In some embodiments, the cell is a cell of stratum opticum. In some embodiments, the cell is a cell of membrana limitans interna. In some embodiments, the cell a retinal pigment epithelium (RPE) cell.

[0051] In some embodiments, described herein is one or more polynucleotides encoding the system described herein. In some embodiments, the one or more polynucleotides comprise a single polynucleotide encoding at least the heterologous polypeptide and the heterologous polynucleotide. In some embodiments, the single polynucleotide further encodes the guide nucleic acid. In some embodiments, the single polynucleotide has a size of less than or equal to about 5 kilobases (kb). In some embodiments, the single polynucleotide has a size of less than or equal to about 4.7 kilobases. In some embodiments, the single polynucleotide has a size of less than or equal to about 0.1 kb to about 10 kb. In some embodiments, the single polynucleotide has a size of less than or equal to about 10 kb to about 9 kb, about 10 kb to about 8 kb, about 10 kb to about 7 kb, about 10 kb to about 6 kb, about 10 kb to about 5 kb, about 10 kb to about 4.7 kb, about 10 kb to about 4 kb, about 10 kb to about 3 kb, about 10 kb to about 2 kb, about 10 kb to about 1 kb, about 10 kb to about 0.1 kb, about 9 kb to about 8 kb, about 9 kb to about 7 kb, about 9 kb to about 6 kb, about 9 kb to about 5 kb, about 9 kb to about 4.7 kb, about 9 kb to about 4 kb, about 9 kb to about 3 kb, about 9 kb to about 2 kb, about 9 kb to about 1 kb, about 9 kb to about 0.1 kb, about 8 kb to about 7 kb, about 8 kb to about 6 kb, about 8 kb to about 5 kb, about 8 kb to about 4.7 kb, about 8 kb to about 4 kb, about 8 kb to about 3 kb, about 8 kb to about 2 kb, about 8 kb to about 1 kb, about 8 kb to about 0.1 kb, about 7 kb to about 6 kb, about 7 kb to about 5 kb, about 7 kb to about 4.7 kb, about 7 kb to about 4 kb, about 7 kb to about 3 kb, about 7 kb to

about 2 kb, about 7 kb to about 1 kb, about 7 kb to about 0.1 kb, about 6 kb to about 5 kb, about 6 kb to about 4.7 kb, about 6 kb to about 4 kb, about 6 kb to about 3 kb, about 6 kb to about 2 kb, about 6 kb to about 1 kb, about 6 kb to about 0.1 kb, about 5 kb to about 4.7 kb, about 5 kb to about 4 kb, about 5 kb to about 3 kb, about 5 kb to about 2 kb, about 5 kb to about 1 kb, about 5 kb to about 0.1 kb, about 4.7 kb to about 4 kb, about 4.7 kb to about 3 kb, about 4.7 kb to about 2 kb, about 4.7 kb to about 1 kb, about 4.7 kb to about 0.1 kb, about 4 kb to about 3 kb, about 4 kb to about 2 kb, about 4 kb to about 1 kb, about 4 kb to about 0.1 kb, about 3 kb to about 2 kb, about 3 kb to about 1 kb, about 3 kb to about 0.1 kb, about 2 kb to about 1 kb, about 2 kb to about 0.1 kb, or about 1 kb to about 0.1 kb. In some embodiments, the single polynucleotide has a size of less than or equal to about 10 kb, about 9 kb, about 8 kb, about 7 kb, about 6 kb, about 5 kb, about 4.7 kb, about 4 kb, about 3 kb, about 2 kb, about 1 kb, or about 0.1 kb. In some embodiments, the single polynucleotide has a size of less than or equal to at least about 10 kb, about 9 kb, about 8 kb, about 7 kb, about 6 kb, about 5 kb, about 4.7 kb, about 4 kb, about 3 kb, about 2 kb, or about 1 kb. In some embodiments, the single polynucleotide has a size of less than or equal to at most about 9 kb, about 8 kb, about 7 kb, about 6 kb, about 5 kb, about 4.7 kb, about 4 kb, about 3 kb, about 2 kb, about 1 kb, or about 0.1 kb.

[0052] In some embodiments, the system comprises a guide nucleic acid or one or more polynucleotides encoding a guide nucleic acid, where the guide nucleic acid targets an endogenous target gene described herein. In some embodiments, the guide nucleic acid can be complexed with an actuator moiety described herein. In some embodiments the guide nucleic acid can direct the actuator moiety to the endogenous target gene in the cell.

[0053] In some embodiments, described herein is a composition comprising any component or any combination of components of the system described herein. In some embodiments, the composition comprises at least one of the heterologous polypeptide described herein. In some embodiments, the compositions comprises at least one of the heterologous polynucleotide described herein. In some embodiments, the composition comprises at least one of the heterologous polypeptide described herein and at least one of the heterologous polynucleotide described herein. In some embodiments, the composition can be further formulated into a pharmaceutical composition. For example, the composition can comprise at least one pharmaceutically acceptable carrier.

[0054] Described herein, in some aspects, is a method comprising: increasing expression level of an endogenous target gene encoding a target protein in a cell, via action of a heterologous polypeptide comprising an actuator moiety, wherein the actuator moiety is for binding the endogenous target gene, and wherein the actuator moiety substantially lacks DNA cleavage

activity; and contacting the cell with a heterologous polynucleotide encoding a non-disease causing variant of the endogenous target gene. In some embodiments, the method comprises determining that the subject has certain condition. In some embodiments, the method comprises selecting for the subject to be treated by the method and the system described herein by determining if the subject harbors a mutant allele or a disease-causing allele of the endogenous target gene. In some embodiments, once the subject is determined to harbor the mutant allele or the disease-causing allele of the endogenous target gene, a system described herein, any component of the system described herein, or any combination of the component of the system described herein can be administered to the subject to treat the disease or condition. In some embodiments, the endogenous target gene comprises a disease causing allele of the target protein. In some embodiments, the endogenous target gene comprises a non-disease causing allele of the endogenous target protein. In some embodiments, the non-disease causing allele is a wild type allele. In some embodiments, the endogenous target gene is a member of Pre-mRNA-processing-splicing factor (PRPF) family. In some embodiments, the PRPF comprises PRPF1, PRPF2, PRPF3, PRPF4, PRPF5, PRPF6, PRPF7, PRPF8, PRPF9, PRPF10, PRPF11, PRPF12, PRPF13, PRPF14, PRPF15, PRPF16, PRPF17, PRPF18, PRPF19, PRPF20, PRPF21, PRPF22, PRPF23, PRPF24, PRPF25, PRPF26, PRPF27, PRPF28, PRPF29, PRPF30, PRPF31, a fragment thereof, or a combination thereof. In some embodiments, the PRPF comprises PRPF31. In some embodiments, the endogenous target gene is PRPF31. In some embodiments, described herein is a method for administering the system described herein to a subject in need thereof. In some embodiments, the method comprises determining whether the subject has or is suspected of having retinitis pigmentosa 11 (RP11). Without wishing to be bound by theory, even when the expression levels of both the non-disease causing allele and the disease causing allele of the PRPF31 are increased, the increased expression level of the non-disease causing allele of the PRPF31 can be sufficient to treat or ameliorate a condition (e.g., retinitis pigmentosa (RP), such as RP11) of a cell or a subject comprising the cell.

[0055] In some embodiments, the method increases the expression of the endogenous target gene (e.g., non-disease causing allele thereof) encoding the target protein by at least about 0.01 fold to about 5,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid). In some embodiments, the method increases the expression of the endogenous target gene encoding the target protein by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about

1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid). In some embodiments, the method increases the expression of the endogenous target gene encoding the target protein by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid). In some embodiments, the method increases the expression of the endogenous target gene encoding the target protein by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid). In some embodiments, the method increases the expression of the endogenous target gene encoding the target protein by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid).

[0056] In some embodiments, the method increases the expression of the endogenous target gene encoding the target protein, where the target protein is a PRPF31 (e.g., a non-disease causing PRPF31 variant). In some embodiments, the method increases the expression of PRPF31 by at least about 0.01 fold to about 5,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid). In some embodiments, the method increases the expression of PRPF31 by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid). In some embodiments, the method increases the expression of PRPF31 by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid). In some embodiments, the method increases the expression of PRPF31 by at

least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid). In some embodiments, the method increases the expression of PRPF31 by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold (e.g., as compared to a control cell lacking the heterologous polypeptide and/or the guide nucleic acid).

[0057] In some embodiments, the method increases the expression of PRPF31 (e.g., a non-disease variant of PRPF31) without increasing expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 5,000 fold. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about

5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold.

[0058] In some embodiments, the method increases the expression of PRPF31 (e.g., a non-disease causing PRPF31 variant) without increasing expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 5,000 fold. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05

fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold. In some embodiments, the method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold.

[0059] In some embodiments, the method increases the expression of PRPF31 (e.g., a non-disease causing PRPF31 variant) without increasing expression of TCF3 Fusion Partner (TFPT). In some embodiments, the method, without increasing the expression of TFPT, increases the

expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 5,000 fold. In some embodiments, the method, without increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold. In some embodiments, the method, without increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold. In some embodiments, the method, without increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold. In some embodiments, the method, without

increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold.

[0060] In some embodiments, the increase in the expression level of the PRPF31 in the (e.g., non-disease causing PRPF31 variant) cell as provided herein effects enhanced (or increased) cilium length and/or incidence, as compared to that in a control cell comprising a mutant allele of the PRPF31 that is in absence of the heterologous polypeptide and/or the guide nucleic acid.

[0061] The increase in the cilium length (e.g., average cilium length) can be at least or up to about 1%, at least or up to about 2%, at least or up to about 5%, at least or up to about 10%, at least or up to about 15%, at least or up to about 20%, at least or up to about 25%, at least or up to about 30%, at least or up to about 40%, at least or up to about 50%, at least or up to about 60%, at least or up to about 70%, at least or up to about 80%, at least or up to about 90%, at least or up to about 100%, at least or up to about 120%, at least or up to about 150%, at least or up to about 200%, at least or up to about 300%, at least or up to about 400%, or at least or up to about 500%, as compared to such control cell. The increase in the cilium length (e.g., average cilium length) can be at least or up to about 0.1-fold, at least or up to about 0.2-fold, at least or up to about 0.5-fold, at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 15-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 150-fold, at least or up to about 200-fold, at least or up to about 300-fold, at least or up to about 400-fold, or at least or up to about 500-fold, as compared to such control cell.

[0062] The increase in the cilium incidence (e.g., average cilium incidence) can be at least or up to about 1%, at least or up to about 2%, at least or up to about 5%, at least or up to about 10%, at least or up to about 15%, at least or up to about 20%, at least or up to about 25%, at least or up to about 30%, at least or up to about 40%, at least or up to about 50%, at least or up to about 60%, at least or up to about 70%, at least or up to about 80%, at least or up to about 90%, at least or up to about 100%, at least or up to about 120%, at least or up to about 150%, at least or up to about 200%, at least or up to about 300%, at least or up to about 400%, or at least or up to about 500%, as compared to such control cell. The increase in the cilium incidence (e.g., average cilium

incidence) can be at least or up to about 0.1-fold, at least or up to about 0.2-fold, at least or up to about 0.5-fold, at least or up to about 1-fold, at least or up to about 2-fold, at least or up to about 3-fold, at least or up to about 4-fold, at least or up to about 5-fold, at least or up to about 6-fold, at least or up to about 7-fold, at least or up to about 8-fold, at least or up to about 9-fold, at least or up to about 10-fold, at least or up to about 15-fold, at least or up to about 20-fold, at least or up to about 30-fold, at least or up to about 40-fold, at least or up to about 50-fold, at least or up to about 60-fold, at least or up to about 70-fold, at least or up to about 80-fold, at least or up to about 90-fold, at least or up to about 100-fold, at least or up to about 150-fold, at least or up to about 200-fold, at least or up to about 300-fold, at least or up to about 400-fold, or at least or up to about 500-fold, as compared to such control cell.

[0063] In some embodiments, the increase in the expression level of the PRPF31 (e.g., non-disease causing PRPF31 variant) in the cell as provided herein effects enhanced localization of a ciliary protein to at least a portion of a cilium, as compared to that in a control cell comprising a mutant allele of the PRPF31 that is in absence of the heterologous polypeptide and/or the guide nucleic acid. For example, the ciliary protein can be IFT88 and the at least a portion of the cilium can be a ciliary tip. In another example, the ciliary protein can be a transition zone (TZ) protein (e.g., CC2D2A, RPGRIP1L, etc.), and the at least the portion of the cilium can be a ciliary axoneme.

[0064] In some embodiments, the heterologous polypeptide (e.g., the dCas-transcriptional effector complexed with a guide RNA) can be sufficient to increase the expression level of the PRPF31 in the cell to render the cell functional. Thus, the systems and the methods of the present disclosure may not and need not require a heterologous PRPF31 and/or a heterologous gene encoding thereof to increase the expression level of the PRPF31 in the cell.

[0065] **Table 4** provides an exemplary list of gRNA (e.g., Cas9 gRNAs) spacer sequence that can bind to an endogenous target gene described herein (e.g., PRPF31). A spacer sequence of gRNA as described herein can comprise a polynucleotide sequence (e.g., a consecutive polynucleotide sequence) that exhibits at least or up to about 50%, at least or up to about 55%, at least or up to about 60%, at least or up to about 65%, at least or up to about 70%, at least or up to about 75%, at least or up to about 80%, at least or up to about 85%, at least or up to about 90%, at least or up to about 91%, at least or up to about 92%, at least or up to about 93%, at least or up to about 94%, at least or up to about 95%, at least or up to about 96%, at least or up to about 97%, at least or up to about 98%, at least or up to about 99%, or substantially about 100% sequence identity to a polynucleotide sequence selected from **Table 4** or a complementary

sequence thereof (e.g., one or more members selected from the group consisting of SEQ ID NOs: 700-785).

[0066] **Table 5** provides an additional exemplary list of gRNA (e.g., Cas gRNAs) spacer sequence that can bind to an endogenous target gene described herein (e.g., PRPF31). A spacer sequence of gRNA as described herein can comprise a polynucleotide sequence (e.g., a consecutive polynucleotide sequence) that exhibits at least or up to about 50%, at least or up to about 55%, at least or up to about 60%, at least or up to about 65%, at least or up to about 70%, at least or up to about 75%, at least or up to about 80%, at least or up to about 85%, at least or up to about 90%, at least or up to about 91%, at least or up to about 92%, at least or up to about 93%, at least or up to about 94%, at least or up to about 95%, at least or up to about 96%, at least or up to about 97%, at least or up to about 98%, at least or up to about 99%, or substantially about 100% sequence identity to a polynucleotide sequence selected from **Table 5** or a complementary sequence thereof (e.g., one or more members selected from the group consisting of SEQ ID NOs: 786-1040). In some cases, the spacer sequence of the guide nucleic acid can target a positive-sense strand (+) of the endogenous target gene. In some cases, the spacer sequence of the guide nucleic acid can target a negative-sense strand (-) of the endogenous target gene.

[0067] The systems (e.g., the heterologous polypeptide and/or a guide nucleic acid) and methods thereof as provided herein can target (e.g., bind) at least one target polynucleotide sequence (e.g., a consecutive polynucleotide sequence) found in the polynucleotide sequence of one or more members in **Table 6**. The at least one target polynucleotide sequence can comprise at least or up to about 1, at least or up to about 2, at least or up to about 3, at least or up to about 4, at least or up to about 5, at least or up to about 6, at least or up to about 7, at least or up to about 8, at least or up to about 9, at least or up to about 10, at least or up to about 15, or at least or up to about 20 target polynucleotide sequence(s). The at least one target polynucleotide sequence can have a length of at least or up to about 6 nucleobases, at least or up to about 8 nucleobases, at least or up to about 10 nucleobases, at least or up to about 12 nucleobases, at least or up to about 16 nucleobases, at least or up to about 18 nucleobases, at least or up to about 20 nucleobases, at least or up to about 22 nucleobases, at least or up to about 24 nucleobases, at least or up to about 26 nucleobases, at least or up to about 28 nucleobases, at least or up to about 30 nucleobases, at least or up to about 32 nucleobases, at least or up to about 34 nucleobases, at least or up to about 36 nucleobases, at least or up to about 38 nucleobases, at least or up to about 40 nucleobases, at least or up to about 45 nucleobases, or at least or up to about 50 nucleobases.

[0068] In some cases, at least a portion of a positive-sense strand (+) of the endogenous target gene can be targeted. The at least the portion of the positive-sense strand can comprise a polynucleotide sequence that exhibits at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or substantially about 100% sequence identity to a consecutive polynucleotide sequence found in SEQ ID NO: 1041 or SEQ ID NO: 1042.

[0069] In some cases, at least a portion of a negative-sense strand (-) of the endogenous target gene can be targeted. The at least the portion of the negative-sense strand can comprise a polynucleotide sequence that exhibits at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or substantially about 100% sequence identity to a consecutive polynucleotide sequence found in SEQ ID NO: 1043 or SEQ ID NO: 1044.

Heterologous polypeptide comprising an actuator moiety

[0070] In various aspects of the present disclosure, the heterologous polypeptide comprising the actuator moiety, as disclosed herein, can be utilized for binding a target gene, such as an endogenous target gene (e.g., a chromosomal DNA sequence). The actuator moiety can be a nuclease, such as an endonuclease (e.g., a heterologous endonuclease). Suitable nucleases include, but are not limited to, CRISPR-associated (Cas) proteins or Cas nucleases including type I CRISPR-associated (Cas) polypeptides, type II CRISPR-associated (Cas) polypeptides, type III CRISPR-associated (Cas) polypeptides, type IV CRISPR-associated (Cas) polypeptides, type V CRISPR-associated (Cas) polypeptides, and type VI CRISPR-associated (Cas) polypeptides; zinc finger nucleases (ZFN); transcription activator-like effector nucleases (TALEN); meganucleases; RNA-binding proteins (RBP); CRISPR-associated RNA binding proteins; recombinases; flippases; transposases; Argonaute (Ago) proteins (e.g., prokaryotic Argonaute (pAgo), archaeal Argonaute (aAgo), and eukaryotic Argonaute (eAgo)); any derivative thereof; any variant thereof and any fragment thereof.

[0071] In some embodiments, the actuator moiety can comprise a DNA nuclease such as an engineered (e.g., programmable or targetable) DNA nuclease that is nuclease-deficient. In some embodiments, the actuator moiety can comprise a nuclease-null DNA binding protein derived from a DNA nuclease that does not induce transcriptional activation or repression of a target

DNA sequence unless it is present in a complex with one or more heterologous gene effectors of the disclosure. In some embodiments, the actuator moiety can comprise a nuclease-null DNA binding protein derived from a DNA nuclease that can induce transcriptional activation or repression of a target DNA sequence (e.g., which can be altered or augmented by the presence of a heterologous gene effector of the disclosure).

[0072] In some embodiments, the actuator moiety can comprise an RNA nuclease such as an engineered (e.g., programmable or targetable) RNA nuclease. In some embodiments, the actuator moiety can comprise a nuclease-null RNA binding protein derived from an RNA nuclease that does not induce transcriptional activation or repression of a target RNA sequence unless it is present in a complex with one or more heterologous gene effectors of the disclosure. In some embodiments, the actuator moiety can comprise a nuclease-null RNA binding protein derived from a RNA nuclease that can induce transcriptional activation or repression of a target RNA sequence (e.g., which can be altered or augmented by the presence of a heterologous gene effector of the disclosure).

[0073] In some embodiments, the actuator moiety can comprise a nucleic acid-guided targeting system. In some embodiments, the actuator moiety can comprise a DNA-guided targeting system. In some embodiments, the actuator moiety can comprise an RNA-guided targeting system. The nucleic acid-guided targeting system can comprise and utilize, for example, a guide nucleic acid sequence that facilitates specific binding of a CRISPR-Cas system (e.g., a nuclease deficient form thereof, such as dCas9 or dCas14) to a target gene (e.g., target endogenous gene) or target gene regulatory sequence. Binding specificity can be determined by use of a guide nucleic acid, such as a single guide RNA (sgRNA) or a part thereof. In some embodiments, the use of different sgRNAs allows the compositions and methods of the disclosure to be used with (e.g., targeted to) different target genes (e.g., target endogenous genes) or target gene regulatory sequences.

[0074] Prokaryotic CRISPR-Cas (Clustered regularly interspaced short palindromic repeats-CRISPR associated) systems, for example, Class II CRISPR-Cas systems such as Cas9 and Cpf1, can be repurposed as a tool for regulation of gene expression, epigenome editing, and chromatin looping in compositions and methods of the disclosure. Nuclease-deactivated Cas (dCas) proteins complexed with heterologous gene effectors can allow for regulation of expression of target genes (e.g., target endogenous genes) adjacent to a site bound by the dCas.

[0075] In some embodiments, the actuator moiety can comprise a CRISPR-associated (Cas) protein or a Cas nuclease that functions in a non-naturally occurring CRISPR (Clustered

Regularly Interspaced Short Palindromic Repeats)/Cas (CRISPR-associated) system. In bacteria, this system can provide adaptive immunity against foreign DNA.

[0076] In a wide variety of organisms including diverse mammals, animals, plants, microbes, and yeast, a CRISPR/Cas system (e.g., modified and/or unmodified) can be utilized as a genome engineering tool, or can be modified to direct specific binding of engineered proteins to target loci as disclosed herein. A CRISPR/Cas system can comprise a guide nucleic acid such as a guide RNA (gRNA) complexed with a Cas protein for targeted regulation of gene expression and/or activity or nucleic acid binding. An RNA-guided Cas protein (e.g., a Cas nuclease such as a Cas9 nuclease) can specifically bind a target polynucleotide (e.g., DNA) in a sequence-dependent manner. The Cas protein, if possessing nuclease activity, can cleave the DNA.

[0077] In some cases, the Cas protein is mutated and/or modified to yield a nuclease deficient protein or a protein with decreased nuclease activity relative to a wild-type Cas protein. A nuclease deficient protein can retain the ability to bind DNA, but may lack or have reduced nucleic acid cleavage activity.

[0078] In some embodiments, the actuator moiety can comprise a Cas protein that forms a complex with a guide nucleic acid, such as a guide RNA or a part thereof. In some embodiments, the actuator moiety can comprise a Cas protein that forms a complex with a single guide nucleic acid, such as a single guide RNA (sgRNA). In some embodiments, the actuator moiety can comprise a RNA-binding protein (RBP) optionally complexed with a guide nucleic acid, such as a guide RNA (e.g., sgRNA), which is able to form a complex with a Cas protein. In some embodiments, the actuator moiety can comprise a nuclease-null DNA binding protein derived from a DNA nuclease that can induce transcriptional activation or repression of a target DNA sequence. In some embodiments, the actuator moiety can comprise a nuclease-null RNA binding protein derived from a RNA.

[0079] A guide nucleic acid used in compositions and methods of the disclosure can comprise a spacer sequence that can bind to an endogenous target gene described herein. The spacer sequence can be, for example, at least 8, at least 9, at least 10, at least 11, at least 12, at least 13, at least 14, at least 15, at least 16, at least 17, at least 18, at least 19, at least 20, at least 21, at least 22, at least 23, at least 24, at least 25, at least 26, at least 27, at least 28, at least 29, at least 30, at least 31, at least 32, at least 33, at least 34, at least 35, at least 36, at least 37, at least 38, at least 39, or at least 40 nucleotides.

[0080] In some embodiments, a spacer sequence of a guide nucleic acid used in compositions and methods of the disclosure is at most at most 10, at most 11, at most 12, at most 13, at most 14, at most 15, at most 16, at most 17, at most 18, at most 19, at most 20, at most 21, at most 22,

at most 23, at most 24, at most 25, at most 26, at most 27, at most 28, at most 29, at most 30, at most 31, at most 32, at most 33, at most 34, at most 35, at most 36, at most 37, at most 38, at most 39, or at most 40 nucleotides.

[0081] In some embodiments, a spacer sequence of a guide nucleic acid used in compositions and methods of the disclosure is between about 8 and about 40 nucleotides, between about 10 and about 40 nucleotides, between about 11 and about 40 nucleotides, between about 12 and about 40 nucleotides, between about 13 and about 40 nucleotides, between about 14 and about 40 nucleotides, between about 15 and about 40 nucleotides, between about 16 and about 40 nucleotides, between about 17 and about 40 nucleotides, between about 18 and about 40 nucleotides, between about 19 and about 40 nucleotides, between about 20 and about 40 nucleotides, between about 22 and about 40 nucleotides, between about 24 and about 40 nucleotides, between about 26 and about 40 nucleotides, between about 28 and about 40 nucleotides, between about 30 and about 40 nucleotides, between about 8 and about 30 nucleotides, between about 10 and about 30 nucleotides, between about 11 and about 30 nucleotides, between about 12 and about 30 nucleotides, between about 13 and about 30 nucleotides, between about 14 and about 30 nucleotides, between about 15 and about 30 nucleotides, between about 16 and about 30 nucleotides, between about 17 and about 30 nucleotides, between about 18 and about 30 nucleotides, between about 19 and about 30 nucleotides, between about 20 and about 30 nucleotides, between about 22 and about 30 nucleotides, between about 24 and about 30 nucleotides, between about 26 and about 30 nucleotides, between about 28 and about 30 nucleotides, between about 8 and about 25 nucleotides, between about 10 and about 25 nucleotides, between about 11 and about 25 nucleotides, between about 12 and about 25 nucleotides, between about 13 and about 25 nucleotides, between about 14 and about 25 nucleotides, between about 15 and about 25 nucleotides, between about 16 and about 25 nucleotides, between about 17 and about 25 nucleotides, between about 18 and about 25 nucleotides, between about 19 and about 25 nucleotides, between about 20 and about 25 nucleotides, between about 22 and about 25 nucleotides, between about 24 and about 25 nucleotides, between about 8 and about 20 nucleotides, between about 10 and about 20 nucleotides, between about 11 and about 20 nucleotides, between about 12 and about 20 nucleotides, between about 13 and about 20 nucleotides, between about 14 and about 20 nucleotides, between about 15 and about 20 nucleotides, between about 16 and about 20 nucleotides, between about 17 and about 20 nucleotides, between about 18 and about 20 nucleotides, between about 19 and about 20 nucleotides, between about 8 and about 18 nucleotides, between about 10 and about 18

nucleotides, between about 11 and about 18 nucleotides, between about 12 and about 18 nucleotides, between about 13 and about 18 nucleotides, between about 14 and about 18 nucleotides, between about 15 and about 18 nucleotides, between about 16 and about 18 nucleotides, between about 8 and about 16 nucleotides, between about 10 and about 16 nucleotides, between about 11 and about 16 nucleotides, between about 12 and about 16 nucleotides, between about 13 and about 16 nucleotides, between about 14 and about 16 nucleotides, or between about 15 and about 16 nucleotides. In some embodiments, a guide nucleic acid can be a guide RNA or a part thereof.

[0082] Non-limiting examples of a guide RNA scaffold sequence are provided in **Table 2**. In some embodiments, the guide RNA scaffold sequence can comprise a polynucleotide sequence (e.g., a consecutive polynucleotide sequence) that exhibits at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or substantially about 100% sequence identity to the polynucleotide sequence of one or more members selected from **Table 2** (e.g., one or more members selected from the group consisting of SEQ ID NOs. 500-596).

[0083] In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of one or more members selected from **Table 2** and (ii) a spacer sequence. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 530, (ii) a spacer sequence, and (iii) the polynucleotide sequence of TT. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 532, (ii) a spacer sequence, and (iii) the polynucleotide sequence of TTTTA. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 534, (ii) a spacer sequence, and (iii) the polynucleotide sequence of TTTTG. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 536, (ii) a spacer sequence, and (iii) the polynucleotide sequence of SEQ ID NO: 537. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 538, (ii) a spacer sequence, and (iii) the polynucleotide sequence of SEQ ID NO: 539. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 541, (ii) a spacer sequence, and (iii) the polynucleotide sequence of SEQ ID NO: 542. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 543, (ii) a spacer sequence, and (iii) the polynucleotide sequence of SEQ ID NO: 544. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 549, (ii) a spacer

sequence, and (ii) the polynucleotide sequence of SEQ ID NO: 550. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 551, (ii) a spacer sequence, and (ii) the polynucleotide sequence of SEQ ID NO: 552. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 554, (ii) a spacer sequence, and (ii) the polynucleotide sequence of TTTTA. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 564, (ii) a spacer sequence, and (ii) the polynucleotide sequence of SEQ ID NO: 550. In some cases, a guide RNA may comprise, from 5' to 3', (i) the polynucleotide sequence of SEQ ID NO: 565, (ii) a spacer sequence, and (ii) the polynucleotide sequence of SEQ ID NO: 550.

[0084] Non-limiting examples of a guide RNA scaffold fragment sequence are provided in **Table 3**. In some embodiments, the guide RNA scaffold sequence can comprise a polynucleotide sequence (e.g., a consecutive polynucleotide sequence) that exhibits at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or substantially about 100% sequence identity to the polynucleotide sequence of one or more members selected from **Table 3** (e.g., one or more members selected from the group consisting of SEQ ID NOs. 597-601).

[0085] Any suitable CRISPR/Cas system can be used. A CRISPR/Cas system can be referred to using a variety of naming systems. A CRISPR/Cas system can be a type I, a type II, a type III, a type IV, a type V, a type VI system, or any other suitable CRISPR/Cas system. A CRISPR/Cas system as used herein can be a Class 1, Class 2, or any other suitably classified CRISPR/Cas system. Class 1 or Class 2 determination can be based upon the genes encoding the effector module. Class 1 systems generally have a multi-subunit crRNA-effector complex, whereas Class 2 systems generally have a single protein, such as Cas9, Cpf1, C2c1, C2c2, C2c3 or a crRNA-effector complex. A Class 1 CRISPR/Cas system can use a complex of multiple Cas proteins to effect regulation. A Class 1 CRISPR/Cas system can comprise, for example, type I (e.g., I, IA, IB, IC, ID, IE, IF, IU), type III (e.g., III, IIIA, IIIB, IIIC, IIID), and type IV (e.g., IV, IVA, IVB) CRISPR/Cas type. A Class 2 CRISPR/Cas system can use a single large Cas protein to effect regulation. A Class 2 CRISPR/Cas systems can comprise, for example, type II (e.g., II, IIA, IIB) and type V CRISPR/Cas type. CRISPR systems can be complementary to each other, and/or can lend functional units in trans to facilitate CRISPR locus targeting.

[0086] When an actuator moiety can comprise a Cas protein or derivative thereof, the Cas protein or derivative thereof can be a Class 1 or a Class 2 Cas protein. A Cas protein can be a type I, type

II, type III, type IV, type V Cas protein, or type VI Cas protein. A Cas protein can comprise one or more domains. Non-limiting examples of domains include, guide nucleic acid recognition and/or binding domain, nuclease domains (e.g., DNase or RNase domains, RuvC, HNH), DNA binding domain, RNA binding domain, helicase domains, protein-protein interaction domains, and dimerization domains. A guide nucleic acid recognition and/or binding domain can interact with a guide nucleic acid. A nuclease domain can comprise catalytic activity for nucleic acid cleavage. A nuclease domain can lack catalytic activity to prevent nucleic acid cleavage. A Cas protein can be a chimeric Cas protein or fragment thereof that is fused to other proteins or polypeptides. A Cas protein can be a chimera of various Cas proteins, for example, comprising domains from different Cas proteins.

[0087] Non-limiting examples of Cas proteins include c2c1, C2c2, c2c3, Cas1, Cas1B, Cas2, Cas3, Cas4, Cas5, Cas5e (CasD), CasH, Cas6e, Cas6f, Cas7, Cas8a, Cas8a1, Cas8a2, Cas8b, Cas8c, Cas9 (Csn1 or Csx12), Cas10, Cas1Od, Cas10, Cas1Od, CasF, CasG, CasH, Cpf1, Csy1, Csy2, Csy3, Cse1 (CasA), Cse2 (CasB), Cse3 (CasE), Cse4 (CasC), Cse1, Cse2, Csa5, Csn2, Csm2, Csm3, Csm4, Csm5, Csm6, Cmr1, Cmr3, Cmr4, Cmr5, Cmr6, Csb1, Csb2, Csb3, Csx17, Csx14, Csx10, Csx16, CsaX, Csx3, Csx1, Csx15, Csf1, Csf2, Csf3, Csf4, Cul966, Cas13a, Cas13b, Cas13c, Cas13d, Cas13X, Cas13Y, Cas14 (e.g., Cas14 variants, such as Cas14a, Cas14b, Cas14c, etc.) and homologs or modified versions thereof.

[0088] A Cas protein or fragment or derivative thereof can be from any suitable organism. Non-limiting examples include *Streptococcus pyogenes*, *Streptococcus thermophilus*, *Streptococcus* sp., *Staphylococcus aureus*, *Nocardiopsis dassonvillei*, *Streptomyces pristinae spiralis*, *Streptomyces viridochromogenes*, *Streptomyces viridochromogenes*, *Streptosporangium roseum*, *Streptosporangium roseum*, *Alicyclobacillus acidocaldarius*, *Bacillus pseudomycoloides*, *Bacillus selenitireducens*, *Exiguobacterium sibiricum*, *Lactobacillus delbrueckii*, *Lactobacillus salivarius*, *Microscilla marina*, *Burkholderiales bacterium*, *Polaromonas naphthalenivorans*, *Polaromonas* sp., *Crocospaera watsonii*, *Cyanotheca* sp., *Microcystis aeruginosa*, *Pseudomonas aeruginosa*, *Synechococcus* sp., *Acetohalobium arabaticum*, *Ammonifex degensii*, *Caldicelulosiruptor beccii*, *Candidatus Desulforudis*, *Clostridium botulinum*, *Clostridium difficile*, *Fingoldia magna*, *Natronaerobius thermophilus*, *Pelotomaculum thermopropionicum*, *Acidithiobacillus caldus*, *Acidithiobacillus ferrooxidans*, *Allochromatium vinosum*, *Marinobacter* sp., *Nitrosococcus halophilus*, *Nitrosococcus watsoni*, *Pseudoalteromonas haloplanktis*, *Ktedonobacter racemifer*, *Methanohalobium evestigatum*, *Anabaena variabilis*, *Nodularia spumigena*, *Nostoc* sp., *Arthrospira maxima*, *Arthrospira platensis*, *Arthrospira* sp., *Lyngbya* sp., *Microcoleus chthonoplastes*, *Oscillatoria* sp., *Petrogla mobilis*, *Thermosiphon africanus*,

Acaryochloris marina, *Leptotrichia shahii*, and *Francisella novicida*. In some aspects, the organism is *Streptococcus pyogenes* (*S. pyogenes*). In some aspects, the organism is *Staphylococcus aureus* (*S. aureus*). In some aspects, the organism is *Streptococcus thermophilus* (*S. thermophilus*).

[0089] A Cas protein can be derived from a variety of bacterial species including, but not limited to, *Veillonella atypical*, *Fusobacterium nucleatum*, *Filifactor alocis*, *Solobacterium moorei*, *Coprococcus catus*, *Treponema denticola*, *Peptoniphilus duerdenii*, *Catenibacterium mitsuokai*, *Streptococcus mutans*, *Listeria innocua*, *Staphylococcus pseudintermedius*, *Acidaminococcus intestine*, *Olsenella uli*, *Oenococcus kitaharae*, *Bifidobacterium bifidum*, *Lactobacillus rhamnosus*, *Lactobacillus gasseri*, *Finegoldia magna*, *Mycoplasma mobile*, *Mycoplasma gallisepticum*, *Mycoplasma ovipneumoniae*, *Mycoplasma canis*, *Mycoplasma synoviae*, *Eubacterium rectale*, *Streptococcus thermophilus*, *Eubacterium dolichum*, *Lactobacillus coryniformis* subsp. *Torquens*, *Ilyobacter polytropus*, *Ruminococcus albus*, *Akkermansia muciniphila*, *Acidothermus cellulolyticus*, *Bifidobacterium longum*, *Bifidobacterium dentium*, *Corynebacterium diphtheria*, *Elusimicrobium minutum*, *Nitratifractor salsuginis*, *Sphaerochaeta globus*, *Fibrobacter succinogenes* subsp. *Succinogenes*, *Bacteroides fragilis*, *Capnocytophaga ochracea*, *Rhodopseudomonas palustris*, *Prevotella micans*, *Prevotella ruminicola*, *Flavobacterium columnare*, *Aminomonas paucivorans*, *Rhodospirillum rubrum*, *Candidatus Puniceispirillum marinum*, *Verminephrobacter eiseniae*, *Ralstonia syzygii*, *Dinoroseobacter shibae*, *Azospirillum*, *Nitrobacter hamburgensis*, *Bradyrhizobium*, *Wolinella succinogenes*, *Campylobacter jejuni* subsp. *Jejuni*, *Helicobacter mustelae*, *Bacillus cereus*, *Acidovorax ebreus*, *Clostridium perfringens*, *Parvibaculum lavamentivorans*, *Roseburia intestinalis*, *Neisseria meningitidis*, *Pasteurella multocida* subsp. *Multocida*, *Sutterella wadsworthensis*, *proteobacterium*, *Legionella pneumophila*, *Parasutterella excrementihominis*, *Wolinella succinogenes*, and *Francisella novicida*.

[0090] A Cas protein as used herein can be a wildtype or a modified form of a Cas protein. A Cas protein can be an active variant, inactive variant, or fragment of a wild type or modified Cas protein. A Cas protein can comprise an amino acid change such as a deletion, insertion, substitution, variant, mutation, fusion, chimera, or any combination thereof relative to a wild-type version of the Cas protein (e.g., a wild-type version of Cas14). A Cas protein can be a polypeptide with at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%,

or 100% sequence identity or sequence similarity to a wild type Cas protein. A Cas protein can be a polypeptide with at most about 5%, at most about 10%, at most about 20%, at most about 30%, at most about 40%, at most about 50%, at most about 60%, at most about 70%, at most about 80%, at most about 90%, or at most about 100% sequence identity and/or sequence similarity to a wild type exemplary Cas protein. Variants or fragments can comprise at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity or sequence similarity to a wild type or modified Cas protein or a portion thereof. Variants or fragments can be targeted to a nucleic acid locus in complex with a guide nucleic acid while lacking nucleic acid cleavage activity.

[0091] A Cas protein can comprise one or more nuclease domains, such as DNase domains. For example, a Cas9 protein can comprise a RuvC-like nuclease domain and/or an HNH-like 20 nuclease domain. The in a nuclease active form of Cas9, RuvC and HNH domains can each cut a different strand of double-stranded DNA to make a double-stranded break in the DNA. A Cas protein can comprise only one nuclease domain (e.g., Cpf1 comprises RuvC domain but lacks HNH domain). In some embodiments, nuclease domains are absent. In some embodiments, nuclease domains are present but inactive or have reduced or minimal activity. In some embodiments, nuclease domains are present and active.

[0092] One or a plurality of the nuclease domains (e.g., RuvC, HNH) of a Cas protein can be deleted or mutated so that they are no longer functional or comprise reduced nuclease activity. For example, in a Cas protein comprising at least two nuclease domains (e.g., Cas9), if one of the nuclease domains is deleted or mutated, the resulting Cas protein, known as a nickase, can generate a single-strand break at a CRISPR RNA (crRNA) recognition sequence within a double-stranded DNA but not a double-strand break. Such a nickase can cleave the complementary strand or the non-complementary strand, but may not cleave both. If all of the nuclease domains of a Cas protein (e.g., both RuvC and HNH nuclease domains in a Cas9 protein; RuvC nuclease domain in a Cpf1 protein) are deleted or mutated, the resulting Cas protein can have a reduced or no ability to cleave both strands of a double-stranded DNA. An example of a mutation that can convert a Cas9 protein into a nickase is a D10A (aspartate to alanine at position 10 of Cas9) mutation in the RuvC domain of Cas9 from *S. pyogenes*. H939A (histidine to alanine at amino acid position 839) or H840A (histidine to alanine at amino acid position 840) in the HNH domain of Cas9 from *S. pyogenes* can convert the Cas9 into a nickase. An example of a mutation that can

convert a Cas9 protein into a dead Cas9 is a D10A (aspartate to alanine at position 10 of Cas9) mutation in the RuvC domain and H939A (histidine to alanine at amino acid position 839) or H840A (histidine to alanine at amino acid position 840) in the HNH domain of Cas9 from *S. pyogenes*.

[0093] A nuclease dead Cas protein can comprise one or more mutations relative to a wild-type version of the protein. The mutation can result in no more than 90%, no more than 80%, no more than 70%, no more than 60%, no more than 50%, no more than 40%, no more than 30%, no more than 20%, no more than 10%, no more than 5%, or no more than 1% of the nucleic acid-cleaving activity in one or more of the plurality of nucleic acid-cleaving domains of the wild-type Cas protein. The mutation can result in one or more of the plurality of nucleic acid-cleaving domains retaining the ability to cleave the complementary strand of the target nucleic acid but reducing its ability to cleave the non-complementary strand of the target nucleic acid. The mutation can result in one or more of the plurality of nucleic acid-cleaving domains retaining the ability to cleave the non-complementary strand of the target nucleic acid but reducing its ability to cleave the complementary strand of the target nucleic acid. The mutation can result in one or more of the plurality of nucleic acid-cleaving domains lacking the ability to cleave the complementary strand and the non-complementary strand of the target nucleic acid. The residues to be mutated in a nuclease domain can correspond to one or more catalytic residues of the nuclease. For example, residues in the wild type exemplary *S. pyogenes* Cas9 polypeptide such as Asp10, His840, Asn854 and Asn856 can be mutated to inactivate one or more of the plurality of nucleic acid-cleaving domains (e.g., nuclease domains). The residues to be mutated in a nuclease domain of a Cas protein can correspond to residues Asp10, His840, Asn854 and Asn856 in the wild type *S. pyogenes* Cas9 polypeptide, for example, as determined by sequence and/or structural alignment.

[0094] A Cas protein can comprise an amino acid sequence having at least about 5%, at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or 100% sequence identity or sequence similarity to a nuclease domain (e.g., RuvC domain, HNH domain) of a wild-type Cas protein.

[0095] A Cas protein, variant or derivative thereof can be modified to enhance regulation of gene expression by compositions and methods of the disclosure, e.g., as part of a complex disclosed herein. A Cas protein can be modified to increase or decrease nucleic acid binding affinity, nucleic acid binding specificity, enzymatic activity, and/or binding to other factors, such as heterodimerization or oligomerization domains and induce ligands. Cas proteins can also be

modified to change any other activity or property of the protein, such as stability. For example, one or more nuclease domains of the Cas protein can be modified, deleted, or inactivated, or a Cas protein can be truncated to remove domains that are not essential for the desired function of the protein or complex. A Cas protein can be modified to modulate (e.g., enhance or reduce) the activity of the Cas protein for regulating gene expression by a complex of the disclosure that comprises a heterologous gene effector.

[0096] For example, a Cas protein can be coupled (e.g., fused, covalently coupled, or non-covalently coupled) to a heterologous gene effector (e.g., an epigenetic modification domain, a transcriptional activation domain, and/or a transcriptional repressor domain). A Cas protein can be coupled (e.g., fused, covalently coupled, or non-covalently coupled) to an oligomerization or dimerization domain as disclosed herein (e.g., a heterodimerization domain). A Cas protein can be coupled (e.g., fused, covalently coupled, or non-covalently coupled) to a heterologous polypeptide that provides increased or decreased stability. A Cas protein can be coupled (e.g., fused, covalently coupled, or non-covalently coupled) to a sequence that can facilitate degradation of the Cas protein or a complex containing the Cas protein, for example, a degron, such as an inducible degron (e.g., auxin inducible).

[0097] A Cas protein can be coupled (e.g., fused, covalently coupled, or non-covalently coupled) to any suitable number of partners, for example, at least one, at least two, at least three, at least four, or at least five, at least six, at least seven, or at least 8 partners. In some embodiments, a Cas protein of the disclosure is coupled (e.g., fused, covalently coupled, or non-covalently coupled) to at most two, at most three, at most four, at most five, at most six, at most seven, at most eight, or at most ten partners. In some embodiments, a Cas protein of the disclosure is coupled (e.g., fused, covalently coupled, or non-covalently coupled) to 1 – 5, 1 – 4, 1 – 3, 1 – 2, 2 – 5, 2 – 4, 2 – 3, 3 – 5, 3 – 4, or 4 – 5 partners. In some embodiments, a Cas protein of the disclosure is coupled (e.g., fused, covalently coupled, or non-covalently coupled) to one partner. In some embodiments, a Cas protein of the disclosure is coupled (e.g., fused, covalently coupled, or non-covalently coupled) to two partners. In some embodiments, a Cas protein of the disclosure is coupled (e.g., fused, covalently coupled, or non-covalently coupled) to three partners. In some embodiments, a Cas protein of the disclosure is coupled (e.g., fused, covalently coupled, or non-covalently coupled) to four partners. In some embodiments, a Cas protein of the disclosure is coupled (e.g., fused, covalently coupled, or non-covalently coupled) to five partners. In some embodiments, a Cas protein of the disclosure is coupled (e.g., fused, covalently coupled, or non-covalently coupled) to six partners.

[0098] A Cas protein can be a fusion protein, e.g., a fusion comprising the Cas protein and one or more of the partners as disclosed herein. The fused domain or heterologous polypeptide can be located at the N-terminus, the C-terminus, or internally within the Cas protein.

[0099] A partner of the Cas protein (e.g., covalently or non-covalently coupled to a dCas protein as disclosed herein) can be a transcriptional effector (e.g., a transcriptional activator or a transcriptional repressor). The transcriptional effector can be heterologous to the cell as provided herein.

[0100] In some embodiments, the Cas protein and the transcriptional effector (e.g., transcriptional activator) can be fused in a single polypeptide sequence. The Cas protein and the transcriptional effector can be fused directly to one another. Alternatively, the Cas protein and the transcriptional effector can be fused via a peptide linker (or an amino acid linker) that is heterologous to the Cas protein and the transcriptional activator. The peptide linker can be derived from a natural polypeptide sequence. Alternatively, the peptide linker can be a synthetic sequence. The peptide linker can have a length of at least or up to about 1 amino acid residue, at least or up to about 2 amino acid residues, at least or up to about 3 amino acid residues, at least or up to about 4 amino acid residues, at least or up to about 5 amino acid residues, at least or up to about 10 amino acid residues, at least or up to about 15 amino acid residues, at least or up to about 20 amino acid residues, at least or up to about 25 amino acid residues, at least or up to about 30 amino acid residues, at least or up to about 35 amino acid residues, at least or up to about 40 amino acid residues, at least or up to about 45 amino acid residues, at least or up to about 50 amino acid residues, at least or up to about 60 amino acid residues, at least or up to about 70 amino acid residues, at least or up to about 80 amino acid residues, at least or up to about 90 amino acid residues, or at least or up to about 100 amino acid residues. In some cases, the peptide linker can be a GS linker.

[0101] The term “GS linker” or “GS linker sequence,” as used interchangeably herein, generally refers to a peptide linker that mainly comprises glycine and serine residues. Particularly, at least or up to about 60%, at least or up to about 65%, at least or up to about 70%, at least or up to about 75%, at least or up to about 80%, at least or up to about 85%, at least or up to about 90%, at least or up to about 95% or substantially about 100% of the amino acid residues in the GS linker sequence can be selected from glycine and serine residues. The GS linker sequence according to the present invention can, for example, comprise from about 1 to about 50 amino acid residues, from about 1 to about 45 amino acid residues, from about 1 to about 40 amino acid residues, from about 1 to about 35 amino acid residues, or from about 1 to about 30 amino

acid residues, in total. In some cases, the GS linker sequence may not comprise about 10, about 5, about 4, about 3, about 2 or about 1 amino acid residue (s) other than glycine or serine.

[0102] In some embodiments, the transcriptional effector can be a histone epigenetic modifier (or a histone modifier). In some cases, the histone epigenetic modifier can modulate histones through methylation (e.g., a histone methylation modifier, such as an amino acid methyltransferase, e.g., KRAB). In some cases, the histone epigenetic modifier can modulate histones through acetylation. In some cases, the histone epigenetic modifier can modulate histones through phosphorylation. In some cases, the histone epigenetic modifier can modulate histones through ADP-ribosylation. In some cases, the histone epigenetic modifier can modulate histones through glycosylation. In some cases, the histone epigenetic modifier can modulate histones through SUMOylation. In some cases, the histone epigenetic modifier can modulate histones through ubiquitination. In some cases, the histone epigenetic modifier can modulate histones by remodeling histone structure, e.g., via an ATP hydrolysis-dependent process.

[0103] In some embodiments, the transcriptional effector can be a gene epigenetic modifier (or a gene modifier). In some cases, a gene modifier can modulate genes through methylation (e.g., a gene methylation modifier, such as a DNA methyltransferase or DNMT). In some cases, a gene modifier can modulate genes through acetylation.

[0104] In some embodiments, the transcriptional effector is from a family of related histone acetyltransferases. Non-limiting examples of histone acetyltransferases include GNAT subfamily, MYST subfamily, p300/CBP subfamily, HAT1 subfamily, GCN5, PCAF, Tip60, MOZ, MORF, MOF, HBO1, p300, CBP, HAT1, ATF-2, SRC1, and TAFII250.

[0105] In some embodiments, the transcriptional effector is from a histone epigenetic modifier (e.g., a histone lysine methyltransferase, a histone lysine demethylase, or a DNA methylase). Non-limiting examples of histone epigenetic modifier include EZH subfamily, Non-SET subfamily, Other SET subfamily, PRDM subfamily, SET1 subfamily, SET2 subfamily, SUV39 subfamily, SYMD subfamily, ASH1L, EHMT1, EHMT2, EZH1, EZH2, MLL, MLL2, MLL3, MLL4, MLL5, NSD1, NSD2, NSD3, PRDM1, PRDM10, PRDM11, PRDM12, PRDM13, PRDM14, PRDM15, PRDM16, PRDM2, PRDM4, PRDM5, PRDM6, PRDM7, PRDM8, PRDM9, SET1, SET1L, SET2L, SETD2, SETD3, SETD4, SETD5, SETD6, SETD7, SETD8, SETDB1, SETDB2, SETMAR, SUV39H1, SUV39H2, SUV420H1, SUV420H2, SYMD1, SYMD2, SYMD3, SYMD4, and SYMD5.

[0106] Examples of proteins (or fragments thereof) that can be used as a fusion partner to 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 and/or

TAL activation domain (e.g., for activity in plants); and histone epigenetic modifier such as SET1A, SET1B, MLL1 to 5, ASH1, SYMD2, NSD1, JHDM2a/b, UTX, JMJD3, GCN5, PCAF, CBP, p300, TAF1, TIP60/PLIP, MOZMYST3, MORFMYST4, SRC1, ACTR, PI 60, CLOCK, Ten-Eleven Translocation (TET) dioxygenase 1 (TET1CD), TET1, DME, DML1, DML2, ROS1, and the like. An additional example of such gene activating modulator is VP64-p65-Rta fusion polypeptide (VPR).

[0107] Examples of proteins (or fragments thereof) that can be used as a fusion partner to decrease transcription include but are not limited to: transcriptional repressors such as the Kruppel 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, RIZ1, and the like; histone lysine demethylases such as JMJD2A/JHDM3A, JMJD2B, JMJD2C/GASC1, JMJD2D, JARID 1 A/RBP2, JARID1B/PLU-1, JARID 1C/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 HhaI DNA m5c-methyltransferase (M.HhaI), DNA methyltransferase 1 (DNMT1), DNA methyltransferase 3a (DNMT3a), DNA methyltransferase 3b (DNMT3b), MET1, DRM3 (plants), ZMET2, CMT1, CMT2 (plants), and the like; and periphery recruitment elements such as Lamin A, Lamin B, and the like.

[0108] A Cas protein can be provided in any form. For example, a Cas protein can be provided in the form of a protein, such as a Cas protein alone or complexed with a guide nucleic acid as a ribonucleoprotein. A Cas protein can be provided in a complex, for example, complexed with a guide nucleic acid and/or one or more heterologous gene effectors of the disclosure. A Cas protein can be provided in the form of a nucleic acid encoding the Cas protein, such as an RNA (e.g., messenger RNA (mRNA)), or DNA. The nucleic acid encoding the Cas protein can be codon optimized for efficient translation into protein in a particular cell or organism.

[0109] Nucleic acids encoding Cas proteins, fragments, or derivatives thereof can be stably integrated in the genome of a cell. Nucleic acids encoding Cas proteins can be operably linked to a promoter, for example, a promoter that is constitutively or inducibly active in the cell. Nucleic acids encoding Cas proteins can be operably linked to a promoter in an expression construct. Expression constructs can include any nucleic acid constructs capable of directing expression of a gene or other nucleic acid sequence of interest (e.g., a Cas gene) and which can transfer such a nucleic acid sequence of interest to a target cell.

[0110] In some embodiments, a Cas protein, variant or derivative thereof is a nuclease dead Cas (dCas) protein. A dead Cas protein can be a protein that lacks nucleic acid cleavage activity.

[0111] A Cas protein can comprise a modified form of a wild type Cas protein. The modified form of the wild type Cas protein can comprise an amino acid change (e.g., deletion, insertion, or substitution) that reduces the nucleic acid-cleaving activity of the Cas protein. For example, the modified form of the Cas protein can have no more than 90%, no more than 80%, no more than 70%, no more than 60%, no more than 50%, no more than 40%, no more than 30%, no more than 20%, no more than 10%, no more than 5%, or no more than 1% of the nucleic acid-cleaving activity of the wild-type Cas protein (e.g., Cas9 from *S. pyogenes*). The modified form of Cas protein can have no substantial nucleic acid-cleaving activity. When a Cas protein is a modified form that has no substantial nucleic acid-cleaving activity, it can be referred to as enzymatically inactive, “deactivated” and/or “dead” (abbreviated by “d”). A dead Cas protein (e.g., dCas, dCas9, dCas14) can bind to a target polynucleotide but may not cleave or minimally cleaves the target polynucleotide. In some aspects, a dead Cas protein is a dead Cas14 protein. In some aspects, a dead Cas protein is a not a dead Cas14 protein.

[0112] A dCas polypeptide (e.g., dCas14 polypeptide) can associate with a single guide RNA (sgRNA) to activate or repress transcription of a target gene (e.g., target endogenous gene), for example, in combination with heterologous gene effector(s) disclosed herein. sgRNAs can be introduced into cells expressing the Cas or variant thereof, as provided herein. In some cases, such cells can contain one or more different sgRNAs that target the same target gene (e.g., target endogenous gene) or target gene regulatory sequence. In other cases, the sgRNAs target different nucleic acids in the cell (e.g., different target genes, different target gene regulatory sequences, or different sequences within the same target gene or target gene regulatory sequence).

[0113] Enzymatically inactive can refer to a nuclease that can bind to a nucleic acid sequence in a polynucleotide in a sequence-specific manner, but will not cleave a target polynucleotide or will cleave it at a substantially reduced frequency. An enzymatically inactive guide moiety can comprise an enzymatically inactive domain (e.g. nuclease domain). Enzymatically inactive can refer to no activity. Enzymatically inactive can refer to substantially no activity. Enzymatically inactive can refer to essentially no activity. Enzymatically inactive can refer to an activity no more than 1%, no more than 2%, no more than 3%, no more than 4%, no more than 5%, no more than 6%, no more than 7%, no more than 8%, no more than 9%, or no more than 10% activity compared to a comparable wild-type activity (e.g., nucleic acid cleaving activity, wild-type Cas9 or wild-type Cas14 activity).

[0114] In some embodiments, the actuator moiety as disclosed herein does not contain a nucleic acid-guided targeting system. For example, the actuator moiety can include proteins that bind to a target gene (e.g., target endogenous gene) or target gene regulatory sequence based on protein structural features, such as certain nucleases disclosed herein.

[0115] In some embodiments, the wild-type Cas protein that the engineered Cas protein is a modification of has a native amino acid sequence with a length of less than 800 amino acids (e.g., Cas14 or a variant thereof). This relatively small size provides several advantages to the provided engineered Cas protein. For example, the small size can allow the Cas protein to be delivered to a host cell, e.g., a cell of a human patient, via a single adeno-associated virus delivery system that would be otherwise incapable of delivering a larger protein. The native amino acid sequence can have a length that is, for example, between 500 amino acids and 700 amino acids, e.g., between 500 amino acids and 620 amino acids, between 540 amino acids and 660 amino acids, between 560 amino acids and 680 amino acids, or between 580 amino acids and 700 amino acids. In terms of upper limits, the native amino acid sequence can have a length that is less than 700 amino acids, e.g., less than 680 amino acids, less than 660 amino acids, less than 640 amino acids, less than 620 amino acids, less than 600 amino acids, less than 580 amino acids, less than 560 amino acids, less than 540 amino acids, or less than 520 amino acids. In terms of lower limits, the native amino acid sequence can have an length that is greater than 500 amino acids, e.g., greater than 520 amino acids, greater than 540 amino acid, greater than 560 amino acids, greater than 580 amino acids, greater than 600 amino acids, greater than 620 amino acids, greater than 640 amino acids, greater than 660 amino acids, or greater than 700 amino acids. Larger lengths, e.g., greater than 700 amino acids, and smaller lengths, e.g., less than 500 amino acids, are also contemplated.

[0116] In some embodiments, the modified amino acid sequence of the engineered Cas protein includes one or more substitutions in the native amino acid sequence, where the positions of at least some of these substitutions follow one or more particular rules determined to have surprising advantages for the characteristics of the engineered Cas protein. For example, the particular substitution rules have been selected for their ability to produce engineered Cas proteins capable of functioning within eukaryotic cells. According to these particular rules, all or some of the one or more substitutions in the native amino acid sequence are either (1) within or no more than 30 amino acids downstream of a (D/E/K/N)X(R/F)(E/K)N motif of the native amino acid sequence, (2) at or no more than 30 amino acids upstream or downstream of position 241 of the native amino acid sequence, (3) at or no more than 30 amino acids upstream or downstream of position 516 of the native amino acid sequence, and/or (4) having an electrically charged amino acid in the native amino acid sequence.

[0117] In some embodiments, the native amino acid sequence includes a (D/E/K/N)X(R/F)(E/K)N motif, and the modified amino acid sequence includes one or more substitutions at positions within or no more than 30 amino acids upstream or downstream of the motif. The modified amino acid sequence can include, for example, one, two, three, four, five, six, seven, eight, nine, ten, or more than ten substitutions within or no more than 30 amino acids upstream or downstream of the motif. At least one of the one or more substitutions to the native amino acid sequence can be, for example, within or no more than 28 amino acids, 26 amino acids, 24 amino acids, 22 amino acids, 20 amino acids, 18 amino acids, 16 amino acids, 14 amino acids, 12 amino acids, or 10 amino acids of the motif. In some embodiments, at least one of the one or more substitutions within or no more than 30 amino acids upstream or downstream of the motif is to an R, A, S, or G. In some embodiments, each of the one or more substitutions within or no more than 30 amino acids upstream or downstream of the motif is independently to an R, A, S, or G. In some embodiments, all of the substitutions to the native amino acid sequence are at positions within or no more than 30 amino acids upstream or downstream of the motif.

[0118] Some embodiments of the present disclosure are directed to a Cas protein that is not a variant of CasX.

[0119] Some embodiments of the present disclosure are directed to small Cas-based regulation of gene expression, such as at the transcriptional and/or translational level. Small Cas proteins can be targeted to DNA and/or RNA, and are much smaller than typical CRISPR effectors, e.g., ranging in size from about 400 amino acids to about 700 amino acids. The small size of can allow such Cas proteins proteins and/or effector domain fusions thereof to be paired with a CRISPR array encoding multiple guide RNAs while remaining under the packaging size limit of various delivery vehicles, such as the versatile adeno-associated virus (AAV) delivery vehicle or non-viral delivery vehicles (e.g., lipid nanoparticles), for primary cell and in vivo delivery.

[0120] In some embodiments, the Cas protein or a variant thereof as provided herein (e.g., a variant of SEQ ID NO: 1 as disclosed herein) can have a size of at most about 800 amino acids, at most about 780 amino acids, at most about 760 amino acids, at most about 750 amino acids, at most about 740 amino acids, at most about 720 amino acids, at most about 700 amino acids, at most about 680 amino acids, at most about 660 amino acids, at most about 650 amino acids, at most about 640 amino acids, at most about 620 amino acids, at most about 600 amino acids, at most about 580 amino acids, at most about 560 amino acids, at most about 550 amino acids, at most about 540 amino acids, at most about 520 amino acids, at most about 500 amino acids, 480 amino acids, at most about 460 amino acids, at most about 450 amino acids, at most about 440 amino acids, at most about 420 amino acids, at most about 400 amino acids, or less.

[0121] Non-limiting examples of Cas protein are provided in **Table 1**. In some embodiments, the Cas protein or the deactivated Cas protein (dCas) as provided herein can comprise a polypeptide sequence (e.g., a consecutive polypeptide sequence) that exhibits at least about 50%, at least about 55%, at least about 60%, at least about 65%, at least about 70%, at least about 75%, at least about 80%, at least about 85%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or substantially about 100% sequence identity to the polypeptide sequence of one or more members selected from **Table 1** (e.g., one or more members selected from the group consisting of SEQ ID NOs. 1-201).

[0122] In some embodiments, the Cas protein or a variant thereof, as provided herein, can comprise the amino acid sequence having at least about 60%, at least about 65%, at least about 70%, at least about 71%, at least about 72%, at least about 73%, at least about 74%, at least about 75%, at least about 76%, at least about 77%, at least about 78%, at least about 79%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater sequence identity to the amino acid sequence of SEQ ID NO: 1. Cas protein or a variant thereof, as provided herein, can comprise the amino acid sequence having at most about 100%, at most about 99%, at most about 98%, at most about 97%, at most about 96%, at most about 95%, at most about 94%, at most about 93%, at most about 92%, at most about 91%, at most about 90%, at most about 89%, at most about 88%, at most about 87%, at most about 86%, at most about 85%, at most about 84%, at most about 83%, at most about 82%, at most about 81%, at most about 80%, at most about 79%, at most about 78%, at most about 77%, at most about 76%, at most about 75%, at most about 74%, at most about 73%, at most about 72%, at most about 71%, at most about 70%, at most about 65%, at most about 60%, or less sequence identity to the amino acid sequence of SEQ ID NO: 1.

[0123] In some embodiments, a Cas protein or a variant thereof as disclosed herein can exhibit a greater cationic charge (e.g., at least 1, at least 2, at least 3, at least 4, at least 5, at least 6, at least 7, at least 8, at least 9, at least 10, or more cationic charges) as compared to the wild-type Cas14. The enhanced cationic charge can (i) enhance complexation of the Cas protein to the guide nucleic acid and/or (ii) enhance complexation of the Cas protein to the target polynucleotide sequence (e.g., endogenous target polynucleotide sequence). In some cases, the Cas protein can comprise one or more substitutions for the enhanced cationic charge. The one or more

substitutions at positions within or no more than 30 amino acids upstream or downstream of the (D/E/K/N)X(R/F)(E/K)N motif of the native amino acid sequence can include, for example, one or more substitutions at positions selected from positions 143, 147, 151, and 154 of the native amino acid sequence. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the one or more substitutions include substitutions are at one or more positions selected from D143, T147, E151, and K154. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the one or more substitutions include one or more substitutions selected from D143R, T147R, E151R, and K154R.

[0124] In some embodiments, the modified amino acid sequence includes one or more substitutions at or no more than 30 amino acids upstream or downstream of position 241 of the native amino acid sequence. The modified amino acid sequence can include, for example, one, two, three, four, five, six, seven, eight, nine, ten, or more than ten substitutions within or no more than 30 amino acids upstream or downstream of position 241. At least one of the one or more substitutions to the native amino acid sequence can be, for example, within or no more than 28 amino acids, 26 amino acids, 24 amino acids, 22 amino acids, 20 amino acids, 18 amino acids, 16 amino acids, 14 amino acids, 12 amino acids, or 10 amino acids of position 241. In some embodiments, at least one of the one or more substitutions within or no more than 30 amino acids upstream or downstream of position 241 is to an R, A, S, or G. In some embodiments, each of the one or more substitutions within or no more than 30 amino acids upstream or downstream of position 241 is independently to an R, A, S, or G. In some embodiments, all of the substitutions to the native amino acid sequence are at positions within or no more than 30 amino acids upstream or downstream of position 241.

[0125] In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the one or more substitutions at positions having an electrically charged amino include substitutions are at one or more positions selected from K11, K73, D143, E151, K154, E241, D318, K330, K457, E425, E462, E507, E527, and E528. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the one or more substitutions include one or more substitutions selected from K11R, K73R, D143R, E151R, K154R, E241R, D318R, K330R, E425N, K457R, E462R, E507R, E527R, and E528R. In some embodiments, the modified amino acid sequence includes a D143R substitution. In some embodiments, the only substitution in the modified amino acid sequence is D143R.

[0126] In some embodiments, the modified amino acid sequence of the engineered Cas protein includes two substitutions in the native amino acid sequence. In some embodiments, the modified amino acid sequence has exactly two substitutions in the native amino acid sequence. In some

embodiments, the modified amino acid sequence includes two substitutions at positions selected from positions 143, 147, 151, 154, 241, 330, 425, 504, 507, 516, 519, 527, and 528. In some embodiments, the modified amino acid sequence has exactly two substitutions, where the exactly two substitutions are at positions selected from positions 143, 147, 151, 154, 241, 330, 425, 504, 507, 516, 519, 527, and 528. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the modified amino acid sequence includes two substitutions at positions selected from D143, T147, E151, K154, E241, K330, E425, N504, E507, N516, N519, E527, and E528. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the modified amino acid sequence has exactly two substitutions, where the exactly two substitutions are at positions selected from D143, T147, E151, K154, E241, K330, E425, N504, E507, N516, N519, E527, and E528.

[0127] In some embodiments, the modified amino acid sequence includes a substitution at position 143 and a substitution at a position selected from positions 147, 151, 154, 241, 330, 425, 504, 507, 516, 519, 527, and 528. In some embodiments, the modified amino acid includes a substitution at position 143 and exactly one other substitution, where the exactly one other substitution is at a position selected from positions 147, 151, 154, 241, 330, 425, 504, 507, 516, 519, 527, and 528. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the modified amino acid sequence includes a substitution at position D143 and a substitution at a position selected from positions T147, E151, K154, E241, K330R, E425N, N504, E507, N516, N519, E527, and E528. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the modified amino acid includes a substitution at position D143 and exactly one other substitution, where the exactly one other substitution is at a position selected from positions T147, E151, K154, E241, K330R, E425N, N504, E507, N516, N519, E527, and E528.

[0128] In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the modified amino acid includes two substitutions selected from D143R, T147R, E151R, E151A, K154R, E241R, N504R, E507R, N516R, N519R, E527R, and E528R. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the modified amino acid includes exactly two substitutions, where the two substitutions are selected from D143R, T147R, E151R, E151A, K154R, E241R, N504R, E507R, N516R, N519R, E527R, and E528R. In some embodiments, e.g., when the native amino acid sequence is the sequence of SEQ ID NO: 1, the modified amino acid includes two substitutions selected from D143R/T147R, D143R/E151R, D143R/E241R, D143R/E425N, D143R/E507R, D143R/N519R, D143R/E527R, D143R/E528R, D143R/R151S, D143R/R151G, and D143R/E151A. In some embodiments, e.g.,

when the native amino acid sequence is the sequence of SEQ ID NO: 1, the modified amino acid includes exactly two substitutions, where the two substitutions are selected from D143R/T147R, D143R/E151R, D143R/E241R, D143R/E425N, D143R/E507R, D143R/N519R, D143R/E527R,, D143R/E528R, D143R/R151S, D143/R151G, and D143R/E151A. In some embodiments, the modified amino acid sequence includes a D143R substitution and a T147R substitution. In some embodiments, the only substitutions in the modified amino acid sequence are a D143R substitution and a T147R substitution.

[0129] In some embodiments, provide herein is a dCas protein or a variant thereof where one or more amino acids of the parental Cas protein from which it is derived have been altered or otherwise removed to reduce or eliminate its nuclease activity. In some embodiments, the amino acids include D326 and D510 with respect to SEQ ID NO: 1. In some embodiments, one or both of D326 and D510 are substituted with an amino acid that reduces, substantially eliminates, or eliminates nuclease activity. In some embodiments, one or both of D326 and D510 are substituted with alanine (e.g., D326A and/or D510A based on SEQ ID NO: 1). In some embodiments, the dCas protein exhibits reduced or eliminated nuclease activity, or nuclease activity is absent or substantially absent within levels of detection.

[0130] In some embodiments, the dCas protein or a variant thereof comprises the amino acid sequence of SEQ ID NO: 1 or a variant thereof having at least about 70%, at least about 71%, at least about 72%, at least about 73%, at least about 74%, at least about 75%, at least about 76%, at least about 77%, at least about 78%, at least about 79%, at least about 80%, at least about 81%, at least about 82%, at least about 83%, at least about 84%, at least about 85%, at least about 86%, at least about 87%, at least about 88%, at least about 89%, at least about 90%, at least about 91%, at least about 92%, at least about 93%, at least about 94%, at least about 95%, at least about 96%, at least about 97%, at least about 98%, at least about 99%, or greater sequence identity to the amino acid sequence of SEQ ID NO: 1.

[0131] In some embodiments, according to any of the Cas protein systems described herein, the target nucleic acid is dsDNA. In such embodiments, dsDNA-targeting specificity is determined, at least in part, by two parameters: the gRNA spacer targeting a protospacer in the target dsDNA (the sequence in the target dsDNA corresponding to the gRNA spacer on the non-complementary DNA strand) and a short sequence, the protospacer-adjacent motif (PAM), located immediately 5' (upstream) of the protospacer on the non-complementary DNA strand. In some embodiments, the PAM is 5'-TTTG-3' or 5'-TTTA-3'. In some embodiments, the PAM is 5'-TTTG-3'. In some embodiments, the PAM is 5'-TTTA-3'.

[0132] In some embodiments, according to any of the Cas protein systems described herein, the target nucleic acid is RNA. In such embodiments, RNA-targeting specificity is determined, at least in part, by the gRNA spacer targeting a protospacer-like sequence in the target RNA (the sequence in the target RNA complementary to the gRNA spacer), and is independent of the sequence located immediately 5' (upstream) of the protospacer-like sequence. In some embodiments, the Cas protein system is also capable of targeting a dsDNA molecule, wherein the gRNA spacer is selected such that it targets a protospacer in the target dsDNA molecule having a PAM selected from 5'-TTTG-3' and 5'-TTTA-3'. In other embodiments, the Cas protein system is incapable of targeting a dsDNA molecule, wherein the gRNA spacer is selected such that any protospacers in the dsDNA molecule targeted by the gRNA spacer do not have a PAM selected from 5'-TTTG-3' and 5'-TTTA-3'.

[0133] In some embodiments, an actuator moiety can comprise a zinc finger nuclease (ZFN) or a variant, fragment, or derivative thereof. ZFN can refer to a fusion between a cleavage domain, such as a cleavage domain of FokI, and at least one zinc finger motif (e.g., at least 2, at least 3, at least 4, or at least 5 zinc finger motifs) which can bind polynucleotides such as DNA and RNA. In some embodiments, a ZFN is used in a targeting moiety of the disclosure to bind a polynucleotide (e.g., target gene or target gene regulatory sequence), but the ZFN does not cleave or substantially does not cleave the polynucleotide, e.g., a nuclease dead ZFN. A ZFN or a variant, fragment, or derivative thereof can be fused to or associated with one of more heterologous gene effectors to form a complex of the disclosure.

[0134] The heterodimerization at certain positions in a polynucleotide of two individual ZFNs in certain orientation and spacing can lead to cleavage of the polynucleotide in nuclease-active ZFN. For example, a ZFN binding to DNA can induce a double-strand break in the DNA. In order to allow two cleavage domains to dimerize and cleave DNA, two individual ZFNs can bind opposite strands of DNA with their C-termini at a certain distance apart. In some cases, linker sequences between the zinc finger domain and the cleavage domain can require the 5' edge of each binding site to be separated by about 5-7 base pairs. In some cases, a cleavage domain is fused to the C-terminus of each zinc finger domain.

[0135] In some embodiments, the cleavage domain of an actuator moiety comprising a ZFN comprises a modified form of a wild type cleavage domain. The modified form of the cleavage domain can comprise an amino acid change (e.g., deletion, insertion, or substitution) that reduces the nucleic acid-cleaving activity of the cleavage domain. For example, the modified form of the cleavage domain can have no more than 90%, no more than 80%, no more than 70%, no more than 60%, no more than 50%, no more than 40%, no more than 30%, no more than 20%, no more

than 10%, no more than 5%, or no more than 1% of the nucleic acid-cleaving activity of the corresponding wild-type cleavage domain. The modified form of the cleavage domain can have no substantial nucleic acid-cleaving activity. In some embodiments, the cleavage domain is enzymatically inactive.

[0136] In some embodiments, an actuator moiety can comprise a “TALEN” or “TAL-effector nuclease” or a variant, fragment, or derivative thereof. TALENs refer to engineered transcription activator-like effector nucleases that generally contain a central domain of DNA-binding tandem repeats and a cleavage domain. TALENs can be produced by fusing a TAL effector DNA binding domain to a DNA cleavage domain. In some cases, a DNA-binding tandem repeat comprises 33-35 amino acids in length and contains two hypervariable amino acid residues at positions 12 and 13 that can recognize at least one specific DNA base pair. A transcription activator-like effector (TALE) protein can be fused to a nuclease such as a wild-type or mutated FokI endonuclease or the catalytic domain of FokI. In some embodiments, a TALEN is used in a targeting moiety of the disclosure to bind a polynucleotide (e.g., target gene or target gene regulatory sequence), but the TALEN does not cleave or substantially does not cleave the polynucleotide, e.g., a nuclease dead TALEN. A TALEN or a variant, fragment, or derivative thereof can be fused to or associated with one of more heterologous gene effectors to form a complex of the disclosure.

[0137] In some embodiments, a TALEN is engineered for reduced nuclease activity. In some embodiments, the nuclease domain of a TALEN comprises a modified form of a wild type nuclease domain. The modified form of the nuclease domain can comprise an amino acid change (e.g., deletion, insertion, or substitution) that reduces the nucleic acid-cleaving activity of the nuclease domain. For example, the modified form of the nuclease domain can have no more than 90%, no more than 80%, no more than 70%, no more than 60%, no more than 50%, no more than 40%, no more than 30%, no more than 20%, no more than 10%, no more than 5%, or no more than 1% of the nucleic acid-cleaving activity of the wild-type nuclease domain. The modified form of the nuclease domain can have no substantial nucleic acid-cleaving activity. In some embodiments, the nuclease domain is enzymatically inactive. A TALEN or a variant, fragment, or derivative thereof can be fused to or associated with one of more heterologous gene effectors to form a complex of the disclosure.

[0138] Several mutations to FokI have been made for its use in TALENs, which, for example, improve cleavage specificity or activity. Such TALENs can be engineered to bind any desired DNA sequence. TALENs can be used to generate gene modifications (e.g., nucleic acid sequence

editing) by creating a double-strand break in a target DNA sequence, which in turn, undergoes NHEJ or HDR.

[0139] A TALE or a variant, fragment, or derivative thereof can be fused to or associated with one of more heterologous gene effectors to form a complex of the disclosure. In some embodiments, the transcription activator-like effector (TALE) protein is fused to a heterologous gene effector and does not comprise a nuclease. In some embodiments, a TALEN does not cleave or substantially does not cleave the polynucleotide, e.g., a nuclease dead TALE. A TALE or a variant, fragment, or derivative thereof can be fused to or associated with one of more heterologous gene effectors to form a complex of the disclosure.

[0140] In some embodiments, the complex of the transcription activator-like effector (TALE) protein and the heterologous gene effector is designed to function as a transcriptional activator. In some embodiments, the complex of the transcription activator-like effector (TALE) protein and the heterologous gene effector is designed to function as a transcriptional repressor. For example, the DNA-binding domain of the transcription activator-like effector (TALE) protein can be fused (e.g., linked) to one or more heterologous gene effectors that comprise transcriptional activation domains, or to one or more heterologous gene effectors that comprise transcriptional repression domains.

[0141] In some embodiments, an actuator moiety can comprise a meganuclease. Meganucleases generally refer to rare-cutting endonucleases or homing endonucleases that can be highly sequence specific. Meganucleases can recognize DNA target sites ranging from at least 12 base pairs in length, e.g., from 12 to 40 base pairs, 12 to 50 base pairs, or 12 to 60 base pairs in length. Meganucleases can be modular DNA-binding nucleases such as any fusion protein comprising at least one catalytic domain of an endonuclease and at least one DNA binding domain or protein specifying a nucleic acid target sequence. The DNA-binding domain can contain at least one motif that recognizes single- or double-stranded DNA. A nuclease-active meganuclease can generate a double-stranded break. In some embodiments, a meganuclease is used in a targeting moiety of the disclosure to bind a polynucleotide (e.g., target gene or target gene regulatory sequence), but the meganuclease does not cleave or substantially does not cleave the polynucleotide, e.g., a nuclease dead meganuclease. A meganuclease or a variant, fragment, or derivative thereof can be fused to or associated with one of more heterologous gene effectors to form a complex of the disclosure.

[0142] The meganuclease can be monomeric or dimeric. In some embodiments, the meganuclease is naturally-occurring (found in nature) or wild-type, and in other instances, the meganuclease is non-natural, artificial, engineered, synthetic, rationally designed, or man-made.

In some embodiments, the meganuclease of the present disclosure includes an I-CreI meganuclease, I-CeuI meganuclease, I-MsoI meganuclease, I-SceI meganuclease, variants thereof, derivatives thereof, and fragments thereof.

[0143] In some embodiments, the nuclease domain of a meganuclease comprises a modified form of a wild type nuclease domain. The modified form of the nuclease domain can comprise an amino acid change (e.g., deletion, insertion, or substitution) that reduces or eliminates the nucleic acid-cleaving activity of the nuclease domain. For example, the modified form of the nuclease domain can have no more than 90%, no more than 80%, no more than 70%, no more than 60%, no more than 50%, no more than 40%, no more than 30%, no more than 20%, no more than 10%, no more than 5%, or no more than 1% of the nucleic acid-cleaving activity of the wild-type nuclease domain. The modified form of the nuclease domain can have no substantial nucleic acid-cleaving activity. In some embodiments, the nuclease domain is enzymatically inactive. In some embodiments, a meganuclease can bind DNA but cannot cleave the DNA. In some embodiments, a nuclease-inactive meganuclease is fused to or associated with one or more heterologous gene effectors to generate a complex of the disclosure.

[0144] In some embodiments, the heterologous polypeptide comprising the actuator moiety (e.g., and/or a complex comprising the heterologous polypeptide) can regulate expression and/or activity of a target gene (e.g., target endogenous gene). In some embodiments, the heterologous polypeptide and/or a complex thereof can edit the sequence of a nucleic acid (e.g., a gene and/or gene product). A nuclease-active Cas protein can edit a nucleic acid sequence by generating a double-stranded break or single-stranded break in a target polynucleotide.

[0145] In some embodiments, the heterologous polypeptide comprising the actuator moiety (e.g., and/or a complex comprising the heterologous polypeptide) can generate a double-strand break in a target polynucleotide, such as DNA. A double-strand break in DNA can result in DNA break repair which allows for the introduction of gene modification(s) (e.g., nucleic acid editing). In some embodiments, a nuclease induces site-specific single-strand DNA breaks or nicks, thus resulting in HDR.

[0146] A double-strand break in DNA can result in DNA break repair which allows for the introduction of gene modification(s) (e.g., nucleic acid editing). DNA break repair can occur via non-homologous end joining (NHEJ) or homology-directed repair (HDR). In HDR, a donor DNA repair template or template polynucleotide that contains homology arms flanking sites of the target DNA can be provided.

[0147] In some embodiments, the heterologous polypeptide comprising the actuator moiety (e.g., and/or a complex comprising the heterologous polypeptide) does not generate a double-strand

break in a target polynucleotide, such as DNA. Binding of the heterologous polypeptide of the complex comprising the heterologous polypeptide (e.g., a complex comprising a dCas-effector and a guide RNA) without a nucleic acid break can be sufficient to regulate expression (e.g., enhance or suppress) of a target gene (e.g., endogenous target gene).

Target gene

[0148] The disclosure provides compositions, methods, and systems for modulating expression of target genes. The target genes can be one or more endogenous target genes, such as a disease causing allele, e.g., a mutant allele. For example, disclosed herein are complexes that comprise a guide moiety and one or more heterologous polypeptides comprising an actuator moiety that can increase or decrease an activity or expression level of a target gene.

[0149] In some embodiments, a target gene or regulatory sequence thereof is endogenous to a cell, for example, present in the cell's genome, or endogenous to a subject, for example, present in the subject's genome. In some embodiments, a target gene or regulatory sequence thereof is not part of an engineered reporter system.

[0150] In some embodiments, a target gene is exogenous to a host subject, for example, a pathogen target gene or an exogenous gene expressed as a result of a therapeutic intervention, such as a gene therapy and/or cell therapy. In some embodiments, a target gene is an exogenous reporter gene. In some embodiments, a target gene is an exogenous synthetic gene.

[0151] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression of a target gene (e.g., upon introducing a complex comprising the heterologous polypeptide into a cell or population of cells). In some embodiments, an expression level is an RNA expression level can be measured by, for example, RNAseq, qPCR, microarray, gene array, FISH, etc. In some embodiments, an expression level is a protein expression level can be measured by, for example, Western Blot, ELISA, multiplex immunoassay, mass spectrometry, NMR, proteomics, flow cytometry, mass cytometry, etc.

[0152] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression of a target gene (e.g., upon introducing a complex comprising the heterologous polypeptide into a cell or population of cells) by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 2-fold, at least about 3 fold, at least about 4 fold, at least about 5 fold, at least about 6 fold, at least about 7 fold, at least about 8 fold, at least about 9 fold, at least about 10 fold, at least about 11 fold, at least about 12 fold, at least about 13 fold, at least about 14, at least fold about 15 fold, at least about 20 fold, at least about 30 fold, at least about 40 fold, at least about 50 fold, at least about 60 fold, at least about 70

fold, at least about 80 fold, at least about 90 fold, at least about 100 fold, at least about 150 fold, at least about 200 fold, at least about 250 fold, at least about 300 fold, at least about 350 fold, at least about 400 fold, at least about 500 fold, at least about 600 fold, at least about 700 fold, at least about 800 fold, at least about 900 fold, at least about 1000 fold, at least about 1500 fold, at least about 2000 fold, or at least about 3000 fold.

[0153] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression of a target gene (e.g., upon introducing a complex comprising the heterologous polypeptide into a cell or population of cells) by at most about 50%, at most about 60%, at most about 70%, at most about 80%, at most about 90%, at most about 2-fold, at most about 3 fold, at most about 4 fold, at most about 5 fold, at most about 6 fold, at most about 7 fold, at most about 8 fold, at most about 9 fold, at most about 10 fold, at most about 11 fold, at most about 12 fold, at most about 13 fold, at most about 14, at most fold about 15 fold, at most about 20 fold, at most about 30 fold, at most about 40 fold, at most about 50 fold, at most about 60 fold, at most about 70 fold, at most about 80 fold, at most about 90 fold, at most about 100 fold, at most about 150 fold, at most about 200 fold, at most about 250 fold, at most about 300 fold, at most about 350 fold, at most about 400 fold, at most about 500 fold, at most about 600 fold, at most about 700 fold, at most about 800 fold, at most about 900 fold, at most about 1000 fold, at most about 1500 fold, at most about 2000 fold, at most about 3000 fold, at most about 5000 fold, or at most about 10000 fold.

[0154] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression of a target gene (e.g., upon introducing a complex comprising the heterologous polypeptide into a cell or population of cells) by about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 2-fold, about 3 fold, about 4 fold, about 5 fold, about 6 fold, about 7 fold, about 8 fold, about 9 fold, about 10 fold, about 11 fold, about 12 fold, about 13 fold, about 14, about 15 fold, about 20 fold, about 30 fold, about 40 fold, about 50 fold, about 60 fold, about 70 fold, about 80 fold, about 90 fold, about 100 fold, about 150 fold, about 200 fold, about 250 fold, about 300 fold, about 350 fold, about 400 fold, about 500 fold, about 600 fold, about 700 fold, about 800 fold, about 900 fold, about 1000 fold, about 1500 fold, about 2000 fold, about 3000 fold, about 5000 fold, or about 10000 fold.

[0155] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression of a target gene (e.g., upon introducing a complex comprising the heterologous polypeptide into a cell or population of cells) from below a limit of detection to a detectable level.

[0156] In some embodiments, the degree in change of expression is relative to before introducing the system of the present disclosure (e.g., a complex comprising the heterologous polypeptide) into the cell or population of cells. In some embodiments, the degree in change of expression is relative to a corresponding control cell or population of cells that are not treated with the system of the present disclosure. In some embodiments, the degree in change of expression is relative to a corresponding control cell or population of cells that are treated with an alternative to the system of the present disclosure.

[0157] In some embodiments, the system and method as disclosed herein can modulate (e.g., increase or decrease) an activity level of a target gene (e.g., upon introducing a complex comprising the heterologous polypeptide into a cell or population of cells). An activity level can be determined by a suitable functional assay for the target gene in question depending on the functional characteristics of the target gene. For example, an activity level of a target gene that is a mitogen could be determined by measuring cell proliferation; an activity level of a target gene that induces apoptosis could be measured by an annexin V assay or other suitable cell death assay; an activity level of an anti-inflammatory cytokine could be measured by an LPS-induced cytokine release assay.

[0158] In some embodiments, the system and method as disclosed herein can modulate (e.g., increase or decrease) the activity of the target gene by at least about 10%, at least about 20%, at least about 30%, at least about 40%, at least about 50%, at least about 60%, at least about 70%, at least about 80%, at least about 90%, at least about 2-fold, at least about 3 fold, at least about 4 fold, at least about 5 fold, at least about 6 fold, at least about 7 fold, at least about 8 fold, at least about 9 fold, at least about 10 fold, at least about 11 fold, at least about 12 fold, at least about 13 fold, at least about 14, at least about 15 fold, at least about 20 fold, at least about 30 fold, at least about 40 fold, at least about 50 fold, at least about 60 fold, at least about 70 fold, at least about 80 fold, at least about 90 fold, at least about 100 fold, at least about 150 fold, at least about 200 fold, at least about 250 fold, at least about 300 fold, at least about 350 fold, at least about 400 fold, at least about 500 fold, at least about 600 fold, at least about 700 fold, at least about 800 fold, at least about 900 fold, at least about 1000 fold, at least about 1500 fold, at least about 2000 fold, or at least about 3000 fold.

[0159] In some embodiments, the system and method as disclosed herein can modulate (e.g., increase or decrease) the activity of the target gene by at most 50%, at most 60%, at most 70%, at most 80%, at most 90%, at most about 2-fold, at most about 3 fold, at most about 4 fold, at most about 5 fold, at most about 6 fold, at most about 7 fold, at most about 8 fold, at most about 9 fold, at most about 10 fold, at most about 11 fold, at most about 12 fold, at most about 13 fold, at most

about 14, at most about 15 fold, at most about 20 fold, at most about 30 fold, at most about 40 fold, at most about 50 fold, at most about 60 fold, at most about 70 fold, at most about 80 fold, at most about 90 fold, at most about 100 fold, at most about 150 fold, at most about 200 fold, at most about 250 fold, at most about 300 fold, at most about 350 fold, at most about 400 fold, at most about 500 fold, at most about 600 fold, at most about 700 fold, at most about 800 fold, at most about 900 fold, at most about 1000 fold, at most about 1500 fold, at most about 2000 fold, at most about 3000 fold, at most about 5000 fold, or at most about 10000 fold.

[0160] In some embodiments, the system to method increases the expression of the endogenous target gene encoding the target protein by at least about 0.01 fold to about 5,000 fold. In some embodiments, the system to method increases the expression of the endogenous target gene encoding the target protein by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold. In some embodiments, the system to method increases the expression of the endogenous

target gene encoding the target protein by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold. In some embodiments, the system to method increases the expression of the endogenous target gene encoding the target protein by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold. In some embodiments, the system to method increases the expression of the endogenous target gene encoding the target protein by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold.

[0161] In some embodiments, the system to method increases the expression of the endogenous target gene encoding the target protein, where the target protein is a PRPF31. In some embodiments, the system to method increases the expression of PRPF31 by at least about 0.01 fold to about 5,000 fold. In some embodiments, the system to method increases the expression of PRPF31 by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about

500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold. In some embodiments, the system to method increases the expression of PRPF31 by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold. In some embodiments, the system to method increases the expression of PRPF31 by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold. In some embodiments, the system to method increases the expression of PRPF31 by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold.

[0162] In some embodiments, the system to method increases the expression of PRPF31 without increasing expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 5,000 fold. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to

about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 or expression of a gene associated with PRPF31, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold.

[0163] In some embodiments, the system to method increases the expression of PRPF31 without increasing expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 5,000 fold. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about

0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold. In some embodiments, the system to method, without increasing the expression of a gene neighboring PRPF31 on chromosome 19 or expression of a gene associated with PRPF31 on chromosome 19, increases the expression of PRPF31 compared to endogenous

expression of PRPF31 by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold.

[0164] In some embodiments, the system to method increases the expression of PRPF31 without increasing expression of TCF3 Fusion Partner (TFPT). In some embodiments, the system to method, without increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 5,000 fold. In some embodiments, the system to method, without increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold to about 0.05 fold, about 0.01 fold to about 0.1 fold, about 0.01 fold to about 0.5 fold, about 0.01 fold to about 1 fold, about 0.01 fold to about 5 fold, about 0.01 fold to about 10 fold, about 0.01 fold to about 50 fold, about 0.01 fold to about 100 fold, about 0.01 fold to about 500 fold, about 0.01 fold to about 1,000 fold, about 0.01 fold to about 5,000 fold, about 0.05 fold to about 0.1 fold, about 0.05 fold to about 0.5 fold, about 0.05 fold to about 1 fold, about 0.05 fold to about 5 fold, about 0.05 fold to about 10 fold, about 0.05 fold to about 50 fold, about 0.05 fold to about 100 fold, about 0.05 fold to about 500 fold, about 0.05 fold to about 1,000 fold, about 0.05 fold to about 5,000 fold, about 0.1 fold to about 0.5 fold, about 0.1 fold to about 1 fold, about 0.1 fold to about 5 fold, about 0.1 fold to about 10 fold, about 0.1 fold to about 50 fold, about 0.1 fold to about 100 fold, about 0.1 fold to about 500 fold, about 0.1 fold to about 1,000 fold, about 0.1 fold to about 5,000 fold, about 0.5 fold to about 1 fold, about 0.5 fold to about 5 fold, about 0.5 fold to about 10 fold, about 0.5 fold to about 50 fold, about 0.5 fold to about 100 fold, about 0.5 fold to about 500 fold, about 0.5 fold to about 1,000 fold, about 0.5 fold to about 5,000 fold, about 1 fold to about 5 fold, about 1 fold to about 10 fold, about 1 fold to about 50 fold, about 1 fold to about 100 fold, about 1 fold to about 500 fold, about 1 fold to about 1,000 fold, about 1 fold to about 5,000 fold, about 5 fold to about 10 fold, about 5 fold to about 50 fold, about 5 fold to about 100 fold, about 5 fold to about 500 fold, about 5 fold to about 1,000 fold, about 5 fold to about 5,000 fold, about 10 fold to about 50 fold, about 10 fold to about 100 fold, about 10 fold to about 500 fold, about 10 fold to about 1,000 fold, about 10 fold to about 5,000 fold, about 50 fold to about 100 fold, about 50 fold to about 500 fold, about 50 fold to about 1,000 fold, about 50 fold to about 5,000 fold, about 100 fold to about 500 fold, about 100 fold to about 1,000 fold, about 100 fold to about 5,000 fold, about 500 fold to about 1,000 fold, about 500 fold to about 5,000 fold, or about 1,000 fold to about 5,000 fold. In some embodiments, the system to method, without increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least about 0.01 fold, about 0.05 fold, about 0.1 fold,

about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold. In some embodiments, the system to method, without increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at least about 0.01 fold, about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, or about 1,000 fold. In some embodiments, the system to method, without increasing the expression of TFPT, increases the expression of PRPF31 compared to endogenous expression of PRPF31 by at least at most about 0.05 fold, about 0.1 fold, about 0.5 fold, about 1 fold, about 5 fold, about 10 fold, about 50 fold, about 100 fold, about 500 fold, about 1,000 fold, or about 5,000 fold.

[0165] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) the activity of the target gene by about 10%, about 20%, about 30%, about 40%, about 50%, about 60%, about 70%, about 80%, about 90%, about 2-fold, about 3 fold, about 4 fold, about 5 fold, about 6 fold, about 7 fold, about 8 fold, about 9 fold, about 10 fold, about 11 fold, about 12 fold, about 13 fold, about 14, about 15 fold, about 20 fold, about 30 fold, about 40 fold, about 50 fold, about 60 fold, about 70 fold, about 80 fold, about 90 fold, about 100 fold, about 150 fold, about 200 fold, about 250 fold, about 300 fold, about 350 fold, about 400 fold, about 500 fold, about 600 fold, about 700 fold, about 800 fold, about 900 fold, about 1000 fold, about 1500 fold, about 2000 fold, about 3000 fold, about 5000 fold, or about 10000 fold.

[0166] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression of a target gene (e.g., upon introducing a complex comprising the heterologous polypeptide into a cell or population of cells) from below a limit of detection to a detectable level.

[0167] In some embodiments, the degree in change of an activity level is relative to before introducing the system of the present disclosure (e.g., a complex comprising the heterologous polypeptide) into the cell or population of cells. In some embodiments, the degree in change of an activity level is relative to a corresponding control cell or population of cells that are not treated with the system of the present disclosure. In some embodiments, the degree in change of an activity level is relative to a corresponding control cell or population of cells that are treated with an alternative to the system of the present disclosure.

[0168] The systems and methods of the present disclosure can, in some cases, elicit changes in expression and/or activity level of a target gene (e.g., target endogenous gene) that persists for longer than can be achieved with alternative compositions and methods (e.g., suppression via

RNAi, e.g., using siRNA). In some embodiments, persistent modulation of gene expression is advantageous as compared to transient modulation.

[0169] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression and/or activity level of a target gene for at least about 1 hour, at least about 2 hours, at least about 3 hours, at least about 4 hours, at least about 5 hours, at least about 6 hours, at least about 7 hours, at least about 8 hours, at least about 9 hours, at least about 10 hours, at least about 12 hours, at least about 14 hours, at least about 18 hours, at least about 20 hours, at least about 1 day, at least about 2 days, at least about 3 days, at least about 4 days, at least about 5 days, at least about 6 days, at least about 7 days, at least about 8 days, at least about 9 days, at least about 10 days, at least about 14 days, at least about 21 days, at least about 28 days, at least about 5 weeks, at least about 6 weeks, at least about 7 weeks, at least about 8 weeks, at least about 9 weeks, at least about 10 weeks, at least about 12 weeks, at least about 14 weeks, at least about 18 weeks, at least about 20 weeks, at least about 26 weeks, or at least about 5 months, at least about 6 months, at least about 9 months, at least about 12 months, or more.

[0170] In some embodiments the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression and/or activity level of a target gene (e.g., target endogenous gene) to above a certain threshold for at most about 1 hour, at most about 2 hours, at most about 3 hours, at most about 4 hours, at most about 5 hours, at most about 6 hours, at most about 7 hours, at most about 8 hours, at most about 9 hours, at most about 10 hours, at most about 12 hours, at most about 14 hours, at most about 18 hours, at most about 20 hours, at most about 1 day, at most about 2 days, at most about 3 days, at most about 4 days, at most about 5 days, at most about 6 days, at most about 7 days, at most about 8 days, at most about 9 days, at most about 10 days, at most about 14 days, at most about 21 days, at most about 28 days, at most about 5 weeks, at most about 6 weeks, at most about 7 weeks, at most about 8 weeks, at most about 9 weeks, at most about 10 weeks, at most about 12 weeks, at most about 14 weeks, at most about 18 weeks, at most about 20 weeks, at most about 26 weeks, or at most about 5 months, at most about 6 months, at most about 9 months, at most about 12 months, or more.

[0171] In some embodiments, the systems and methods as disclosed herein can modulate (e.g., increase or decrease) expression and/or activity level of a target gene (e.g., target endogenous gene) to above a certain threshold for about 1 hour, about 2 hours, about 3 hours, about 4 hours, about 5 hours, about 6 hours, about 7 hours, about 8 hours, about 9 hours, about 10 hours, about 12 hours, about 14 hours, about 18 hours, about 20 hours, about 1 day, about 2 days, about 3 days, about 4 days, about 5 days, about 6 days, about 7 days, about 8 days, about 9 days, about 10 days, about 14 days, about 21 days, about 28 days, about 5 weeks, about 6 weeks, about 7

weeks, about 8 weeks, about 9 weeks, about 10 weeks, about 12 weeks, about 14 weeks, about 18 weeks, about 20 weeks, about 26 weeks, about 5 months, about 6 months, about 9 months, or about 12 months.

[0172] In some embodiments, the target gene (e.g., endogenous target gene) can be a disease-causing allele, such as a mutant variant of a wild type allele. The disease can be a genetic disease, such as a hereditary disorder. Non-limiting examples of the genetic disorder can include Duchenne muscular dystrophy (DMD), hemophilia, cystic fibrosis, Huntington's chorea, familial hypercholesterolemia (LDL receptor defect), hepatoblastoma, Wilson's disease, congenital hepatic porphyria, inherited disorders of hepatic metabolism, Lesch Nyhan syndrome, sickle cell anemia, thalassaemias, xeroderma pigmentosum, Fanconi's anemia, retinitis pigmentosa, ataxia telangiectasia, Bloom's syndrome, retinoblastoma, and Tay-Sachs disease. In some cases, the target gene can be a gene encoding a protein. In some cases, the target gene can be a gene regulatory sequence (e.g., promoters, enhancers, repressors, silencers, insulators, cis-regulatory elements, trans-regulatory elements, epigenetic modification (e.g., DNA methylation) sites, etc.) that can influence expression of a gene encoding a protein of interest as provided herein. For example, target gene regulatory sequences can be physically located outside of the transcriptional unit or open reading frame that encodes a product of the target gene.

[0173] In some embodiments, a target gene regulatory sequence does not contain a nucleotide sequence that is exogenous to the subject or host cell. In some embodiments, a target gene regulatory sequence does not contain an engineered or artificially generated or introduced nucleotide sequence.

[0174] In some embodiments, a target gene (e.g., target endogenous gene) is a gene that is over-expressed or under-expressed in a disease or condition. In some embodiments, a target gene is a gene that is over-expressed or under-expressed in a heritable genetic disease. In some embodiments, the disease is retinitis pigmentosa 11 (RP11).

Heterologous polynucleotide

[0175] In some embodiments, a target gene (e.g., an endogenous target gene) can be a disease causing gene (e.g., a mutant allele), and the systems and compositions of the present disclosure can further comprise a heterologous polynucleotide encoding a non-disease causing gene thereof (e.g., a wild type allele), e.g., as a gene replacement therapy. Accordingly, the methods as disclosed herein can comprise introducing such system or compositions to a cell or to a subject, e.g., contacting the cell with such systems or compositions (e.g., via delivery or expression of such systems or compositions in the cell).

[0176] Thus, the systems and compositions can comprise the non-disease causing wild type or variant of the target gene, as abovementioned. Alternatively or in addition to, the systems and compositions can comprise a heterologous polynucleotide sequence encoding (or comprising) at least the non-disease causing wild type or variant of the target gene (e.g., that of the endogenous target gene) as disclosed herein.

Composition

[0177] In some aspects, the present disclosure provides a composition comprising at least a portion of the system as described, e.g., (i) the heterologous polypeptide comprising the actuator moiety or a heterologous polynucleotide encoding the heterologous polypeptide. (ii) the guide nucleic acid or a heterologous polynucleotide encoding the guide nucleic acid, as disclosed herein, (iii) the heterologous polynucleotide encoding a non-disease causing allele of a gene, for use in any of the methods as disclosed herein. The subject composition can be usable for modifying a cell in vitro, ex vivo, or in vivo. The subject composition can be usable for treating or enhancing a condition of a subject, as disclosed herein.

[0178] The composition as disclosed herein can comprise an active ingredient (e.g., the heterologous polypeptide comprising the actuator moiety, the guide nucleic acid, the heterologous polynucleotide encoding the non-disease causing allele of a gene, etc.) and optionally an additional ingredient (e.g., excipient). If necessary and/or desirable, the composition can be divided, shaped and/or packaged into a desired single- or multi-dose unit or single-or multi-implantation unit.

[0179] In some embodiments, the composition can comprise one or more heterologous polynucleotides encoding the active ingredients as disclosed herein. When there are different members within the active ingredients, each member can be encoded by a different heterologous polynucleotide. Alternatively, two or more (e.g., all of) the ingredients can be encoded by a single heterologous polynucleotide. In some cases, a single heterologous polynucleotide can encode (i) the heterologous polypeptide comprising the actuator moiety (e.g., dCas-transcriptional effector fusion protein, such as dCas-KRAB or dCas-DNMT) and (ii) one or more guide nucleic acids (e.g., at least 1, at least 2, at least 3, at least 4, at least 5, or more guide nucleic acids) for targeting specific region(s) or sequence(s) of the target gene. In some cases, a single heterologous polynucleotide can encode (i) the heterologous polypeptide comprising the actuator moiety (e.g., dCas-transcriptional effector fusion protein, such as dCas-KRAB, dCas-DNMT, or dCas-VPR), (ii) one or more guide nucleic acids (e.g., at least 1, at least 2, at least 3, at least 4, at least 5, or more guide nucleic acids) for targeting specific region(s) or sequence(s) of the target gene, and (iii) the heterologous polynucleotide encoding a non-disease causing allele of

a gene.

[0180] The one or more heterologous polynucleotides can further comprise one or more promoters (or one or more transcriptional control elements, as used interchangeably herein). Different active ingredients encoded by the one or more heterologous polynucleotides can be under the control of the same promoter or different promoters. A promoter as disclosed herein can be active in a eukaryotic, mammalian, non-human mammalian or human cell. The promoter can be an inducible or constitutively active promoter. Alternatively or additionally, the promoter can be tissue or cell specific. Non-limiting examples of suitable eukaryotic promoters (i.e. promoters functional in a eukaryotic cell) can include those from cytomegalovirus (CMV) immediate early, herpes simplex virus (HSV) thymidine kinase, early and late SV40, long terminal repeats (LTRs) from retrovirus, human elongation factor-1 promoter (EF1), a hybrid construct comprising the cytomegalovirus (CMV) enhancer fused to the chicken beta-actin promoter (CAG), murine stem cell virus promoter (MSCV), phosphoglycerate kinase-1 locus promoter (PGK) and mouse metallothionein-I. The promoter can be a fungi promoter. The promoter can be a plant promoter. A database of plant promoters can be found (e.g., PlantProm). 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. In some cases, a promoter as disclosed herein can be a promoter specific for any of the tissues provided herein, or a promoter specific for any of the cell types provided herein.

[0181] A heterologous polynucleotide of the one or more heterologous polynucleotides (e.g., the single heterologous polynucleotide) can have a size of at least or up to about 2.5 kilobases, at least or up to about 2.6 kilobases, at least or up to about 2.7 kilobases, at least or up to about 2.8 kilobases, at least or up to about 2.9 kilobases, at least or up to about 3.0 kilobases, at least or up to about 3.1 kilobases, at least or up to about 3.2 kilobases, at least or up to about 3.3 kilobases, at least or up to about 3.4 kilobases, at least or up to about 3.5 kilobases, at least or up to about 3.6 kilobases, at least or up to about 3.7 kilobases, at least or up to about 3.8 kilobases, at least or up to about 3.9 kilobases, at least or up to about 4.0 kilobases, at least or up to about 4.1 kilobases, at least or up to about 4.2 kilobases, at least or up to about 4.3 kilobases, at least or up to about 4.4 kilobases, at least or up to about 4.5 kilobases, at least or up to about 4.6 kilobases, at least or up to about 4.7 kilobases, at least or up to about 4.8 kilobases, at least or up to about 4.9 kilobases, at least or up to about 5.0 kilobases, at least or up to about 5.5 kilobases, at least or up to about 6.0 kilobases, at least or up to about 6.5 kilobases, at least or up to about 7.0 kilobases, at least or up to about 7.5 kilobases, at least or up to about 8.0 kilobases, at least or up

to about 9.0 kilobases, or at least or up to about 10 kilobases. In some cases, the heterologous polynucleotide of the one or more heterologous polynucleotides (e.g., the single heterologous polynucleotide) can have a size of between about 3 kilobases and about 5 kilobases, between about 3 kilobases and about 4.8 kilobases, between about 3 kilobases and about 4.6 kilobases, between about 3 kilobases and about 4.4 kilobases, between about 3 kilobases and about 4.2 kilobases, between about 3 kilobases and about 4.0 kilobases, between about 3 kilobases and about 3.5 kilobases, between about 3.5 kilobases and about 5 kilobases, between about 3.5 kilobases and about 4.8 kilobases, between about 3.5 kilobases and about 4.6 kilobases, between about 3.5 kilobases and about 4.4 kilobases, between about 3.5 kilobases and about 4.2 kilobases, between about 3.5 kilobases and about 4 kilobases, between about 4 kilobases and about 5 kilobases, between about 4 kilobases and about 4.9 kilobases, between about 4 kilobases and about 4.8 kilobases, between about 4 kilobases and about 4.7 kilobases, between about 4 kilobases and about 4.6 kilobases, between about 4 kilobases and about 4.5 kilobases, between about 4 kilobases and about 4.4 kilobases, between about 4 kilobases and about 4.3 kilobases, between about 4 kilobases and about 4.2 kilobases, or between about 4 kilobases and about 4.1 kilobases.

[0182] A method of delivery of the one or more heterologous polynucleotides provided herein to the cell can involve viral delivery methods or non-viral delivery methods. Thus, the one or more heterologous polynucleotides can be one or more viral vectors (e.g., one or more AAV vectors). Alternatively, the one or more heterologous polynucleotides can be non-viral vectors that are complexed with or encapsulated by non-viral delivery moieties, such as cationic lipids and/or lipid particles (e.g., lipid nanoparticles (LNP)).

[0183] Methods of non-viral delivery of nucleic acids can include lipofection, nucleofection, microinjection, biolistics, virosomes, liposomes, immunoliposomes, polycation or lipid:nucleic acid conjugates, naked DNA, artificial virions, and agent-enhanced uptake of DNA. Cationic and neutral lipids that are suitable for efficient receptor-recognition lipofection of polynucleotides can be used. Delivery can be to cells (e.g. in vitro or ex vivo administration) or target tissues (e.g. in vivo administration).

[0184] In some embodiments, the compositions and systems provided herein are delivered to a subject using a viral vector. In some cases, the viral vector is an adeno-associated viral (AAV) vector. The term “AAV” is an abbreviation for adeno-associated virus, and may be used to refer to the virus itself or a derivative thereof. The term covers all serotypes, subtypes, and both naturally occurring and recombinant forms, except where required otherwise. The abbreviation “rAAV” refers to recombinant adeno-associated virus, also referred to as a recombinant AAV

vector (or “rAAV vector”). The term “AAV” includes AAV1, AAV2, AAV3, AAV4, AAV5, AAV6, AAV7, AAV8, AAV9, AAV10, AAV11, AAV12, rh10, and hybrids thereof, avian AAV, bovine AAV, canine AAV, equine AAV, primate AAV, non-primate AAV, and ovine AAV. The genomic sequences of various serotypes of AAV, as well as the sequences of the native terminal repeats (TRs), Rep proteins, and capsid subunits are known in the art. Such sequences may be found in the literature or in public databases such as GenBank. An “rAAV vector” as used herein refers to an AAV vector comprising a polynucleotide sequence not of AAV origin (i.e., a polynucleotide heterologous to AAV), typically a sequence of interest for the genetic transformation of a cell. In general, the heterologous polynucleotide is flanked by at least one, and generally by two, AAV inverted terminal repeat sequences (ITRs). The term rAAV vector encompasses both rAAV vector particles and rAAV vector plasmids. An rAAV vector may either be single-stranded (ssAAV) or self-complementary (scAAV). An “AAV virus” or “AAV viral particle” or “rAAV vector particle” refers to a viral particle composed of at least one AAV capsid protein and an encapsidated polynucleotide rAAV vector. If the particle comprises a heterologous polynucleotide (i.e., a polynucleotide other than a wild-type AAV genome such as a transgene to be delivered to a mammalian cell), it is typically referred to as an “rAAV vector particle” or simply an “rAAV vector”. Thus, production of rAAV particle necessarily includes production of rAAV vector, as such a vector is contained within an rAAV particle. In some cases, the AAV vector is selected based on the tropism of viral vector. In some embodiments, an AAV vector with tropism for the target tissue (e.g., eye) may be used (e.g., AAV7, AAV8, AAV9) to deliver polynucleotides encoding the compositions and systems provided herein to the target tissue (e.g., eye).

[0185] RNA or DNA viral based systems can be used to target specific cells in the body and trafficking the viral payload to the nucleus of the cell. Viral vectors can be administered directly (in vivo) or they can be used to treat cells in vitro, and the modified cells can optionally be administered (ex vivo). Viral based systems can include retroviral, lentivirus, adenoviral, adeno-associated and herpes simplex virus vectors for gene transfer. Integration in the host genome can occur with the retrovirus, lentivirus, and adeno-associated virus gene transfer methods, which can result in long term expression of the inserted transgene. High transduction efficiencies can be observed in many different cell types and target tissues.

[0186] The tropism of a retrovirus can be altered by incorporating foreign envelope proteins, expanding the potential target population of target cells. Lentiviral vectors are retroviral vectors that can transduce or infect non-dividing cells and produce high viral titers. Selection of a retroviral gene transfer system can depend on the target tissue. Retroviral vectors can comprise

cis-acting long terminal repeats with packaging capacity for up to 6-10 kb of foreign sequence. The minimum cis-acting LTRs can be sufficient for replication and packaging of the vectors, which can be used to integrate the therapeutic gene into the target cell to provide permanent transgene expression. Retroviral vectors can include those based upon murine leukemia virus (MuLV), gibbon ape leukemia virus (GaLV), Simian Immuno deficiency virus (SIV), human immuno deficiency virus (HIV), and combinations thereof.

[0187] An adenoviral-based systems can be used. Adenoviral-based systems can lead to transient expression of the transgene. Adenoviral based vectors can have high transduction efficiency in cells and may not require cell division. High titer and levels of expression can be obtained with adenoviral based vectors. Adeno-associated virus ("AAV") vectors can be used to transduce cells with target nucleic acids, e.g., in the in vitro production of nucleic acids and peptides, and for in vivo and ex vivo gene therapy procedures.

[0188] Packaging cells can be used to form virus particles capable of infecting a host cell. Such cells can include 293 cells, (e.g., for packaging adenovirus), and Psi2 cells or PA317 cells (e.g., for packaging retrovirus). Viral vectors can be generated by producing a cell line that packages a nucleic acid vector into a viral particle. The vectors can contain the minimal viral sequences required for packaging and subsequent integration into a host. The vectors can contain other viral sequences being replaced by an expression cassette for the polynucleotide(s) to be expressed. The missing viral functions can be supplied in trans by the packaging cell line. For example, AAV vectors can comprise ITR sequences from the AAV genome which are required for packaging and integration into the host genome. Viral DNA can be packaged in a cell line, which can contain a helper plasmid encoding the other AAV genes, namely rep and cap, while lacking ITR sequences. The cell line can also be infected with adenovirus as a helper. The helper virus can promote replication of the AAV vector and expression of AAV genes from the helper plasmid. Contamination with adenovirus can be reduced by, e.g., heat treatment to which adenovirus is more sensitive than AAV.

[0189] A host cell can be transiently or non-transiently transfected with one or more vectors described herein. A cell can be transfected as it naturally occurs in a subject. A cell can be taken or derived from a subject and transfected. A cell can be derived from cells taken from a subject, such as a cell line. In some embodiments, a cell transfected with one or more vectors described herein is used to establish a new cell line comprising one or more vector-derived sequences. In some embodiments, a cell transiently transfected with the compositions of the disclosure (such as by transient transfection of one or more vectors, or transfection with RNA), and modified through the activity of an actuator moiety such as a CRISPR complex, is used to establish a

new cell line comprising cells containing the modification but lacking any other exogenous sequence.

[0190] Any suitable vector compatible with the host cell can be used with the methods of the disclosure. Non-limiting examples of vectors for eukaryotic host cells include pXT1, pSG5 (Stratagene™), pSVK3, pBPV, pMSG, and pSVLSV40 (Pharmacia™).

[0191] In some embodiments, the additional ingredient of the composition as disclosed herein can comprise an excipient. Non-limiting examples of the excipient can include solvents, dispersion media, diluents, or other liquid vehicles, dispersion or suspension aids, surface active agents, isotonic agents, thickening or emulsifying agents, preservatives, lipidoids, liposomes, lipid nanoparticles, polymers, lipoplexes, core-shell nanoparticles, peptides, proteins, hyaluronidase, nanoparticle mimics, inert diluents, buffering agents, lubricating agents, oils, and combinations thereof. In some examples, the composition as disclosed herein can include one or more excipients, each in an amount that together increases the stability of (i) the heterologous polypeptide or the heterologous gene encoding thereof and/or (ii) cells or modified cells.

[0192] In some aspects, the present disclosure provides a kit comprising such composition and instructions directing (i) contacting the cell with the composition (e.g., in vitro, ex vivo, or in vivo), or (ii) administration of cells comprising any one of the compositions disclosed herein to a subject. The subject may have or may be suspected of having a condition, such as a hereditary disease.

[0193] In some embodiments, any of the compositions as disclosed herein, can be administered to the subject via orally, intraperitoneally, intravenously, intraarterially, transdermally, intramuscularly, liposomally, via local delivery by catheter or stent, subcutaneously, intraadiposally, or intrathecally. In particular aspects, the compositions and systems provided herein (including polynucleotides encoding said compositions and systems, e.g., contained in an AAV vector) can be administered to a subject via intravenous administration.

[0194] The compositions (e.g., pharmaceutical compositions) as disclosed herein can be suitable for administration to humans. In addition, such compositions can be suitable for administration to any other animal, e.g., to non-human animals, e.g. non-human mammals. Modification of pharmaceutical compositions suitable for administration to humans in order to render the compositions suitable for administration to various animals is well understood, and the ordinarily skilled veterinary pharmacologist can design and/or perform such modification with merely ordinary, if any, experimentation. Subjects to which administration of the pharmaceutical compositions is contemplated include, but are not limited to, humans and/or other primates; mammals, including commercially relevant mammals such as cattle, pigs, horses, sheep, cats,

dogs, mice, and/or rats; and/or birds, including commercially relevant birds such as poultry, chickens, ducks, geese, and/or turkeys.

Cells

[0195] In some embodiments, a cell as provided herein may be referred to as a target cell. In some embodiments, the systems, compositions, and methods as provided herein can be applied to modify a target cell (e.g., modify expression profile of a target gene of the target cell). A target cell can include a wide variety of cell types. A target cell can be in vitro. A target cell can be in vivo. A target cell can be ex vivo. A target cell can be an isolated cell. A target cell can be a cell inside of an organism. A target cell can be an organism. A target cell can be a cell in a cell culture. A target cell can be one of a collection of cells. A target cell can be a mammalian cell or derived from a mammalian cell. A target cell can be a rodent cell or derived from a rodent cell. A target cell can be a human cell or derived from a human cell. A target cell can be a prokaryotic cell or derived from a prokaryotic cell. A target cell can be a bacterial cell or can be derived from a bacterial cell. A target cell can be an archaeal cell or derived from an archaeal cell. A target cell can be a eukaryotic cell or derived from a eukaryotic cell. A target cell can be a pluripotent stem cell. A target cell can be a plant cell or derived from a plant cell. A target cell can be an animal cell or derived from an animal cell. A target cell can be an invertebrate cell or derived from an invertebrate cell. A target cell can be a vertebrate cell or derived from a vertebrate cell. A target cell can be a microbe cell or derived from a microbe cell. A target cell can be a fungi cell or derived from a fungi cell. A target cell can be from a specific organ or tissue.

[0196] A target cell can be a stem cell or progenitor cell. Target cells can include stem cells (e.g., adult stem cells, embryonic stem cells, induced pluripotent stem (iPS) cells) and progenitor cells (e.g., cardiac progenitor cells, neural progenitor cells, etc.). Target cells can include mammalian stem cells and progenitor cells, including rodent stem cells, rodent progenitor cells, human stem cells, human progenitor cells, etc. Clonal cells can comprise the progeny of a cell. A target cell can comprise a target nucleic acid. A target cell can be in a living organism. A target cell can be a genetically modified cell. A target cell can be a host cell.

[0197] A target cell can be a primary cell. For example, cultures of primary cells can be passaged 0 times, 1 time, 2 times, 4 times, 5 times, 10 times, 15 times or more. Cells can be unicellular organisms. Cells can be grown in culture.

[0198] A target cell can be a diseased cell. A diseased cell can have altered metabolic, gene expression, and/or morphologic features. A diseased cell can be a cancer cell, a diabetic cell, and a apoptotic cell. A diseased cell can be a cell from a diseased subject. Exemplary diseases can include blood disorders, cancers, metabolic disorders, eye disorders, organ disorders,

musculoskeletal disorders, cardiac disease, and the like.

[0199] If the target cells are primary cells, they may be harvested from an individual by any method. For example, leukocytes may be harvested by apheresis, leukocytapheresis, density gradient separation, etc. Cells from tissues such as skin, muscle, bone marrow, spleen, liver, pancreas, lung, intestine, stomach, etc. can be harvested by biopsy.

[0200] Non-limiting examples of cells which can be target cells include, but are not limited to, lymphoid cells, such as B cell, T cell (Cytotoxic T cell, Natural Killer T cell, Regulatory T cell, T helper cell), Natural killer cell, cytokine induced killer (CIK) cells; myeloid cells, such as granulocytes (Basophil granulocyte, Eosinophil granulocyte, Neutrophil granulocyte/Hypersegmented neutrophil), Monocyte/Macrophage, Red blood cell (Reticulocyte), Mast cell, Thrombocyte/Megakaryocyte, Dendritic cell; cells from the endocrine system, including thyroid (Thyroid epithelial cell, Parafollicular cell), parathyroid (Parathyroid chief cell, Oxyphil cell), adrenal (Chromaffin cell), pineal (Pinealocyte) cells; cells of the nervous system, including glial cells (Astrocyte, Microglia), Magnocellular neurosecretory cell, Stellate cell, Boettcher cell, and pituitary (Gonadotrope, Corticotrope, Thyrotrope, Somatotrope, Lactotroph); cells of the Respiratory system, including Pneumocyte (Type I pneumocyte, Type II pneumocyte), Clara cell, Goblet cell, Dust cell; cells of the circulatory system, including Myocardiocyte, Pericyte; cells of the digestive system, including stomach (Gastric chief cell, Parietal cell), Goblet cell, Paneth cell, G cells, D cells, ECL cells, I cells, K cells, S cells; enteroendocrine cells, including enterochromaffin cell, APUD cell, liver (Hepatocyte, Kupffer cell), Cartilage/bone/muscle; bone cells, including Osteoblast, Osteocyte, Osteoclast, teeth (Cementoblast, Ameloblast); cartilage cells, including Chondroblast, Chondrocyte; skin cells, including Trichocyte, Keratinocyte, Melanocyte (Nevus cell); muscle cells, including Myocyte; urinary system cells, including Podocyte, Juxtaglomerular cell, Intraglomerular mesangial cell/Extraglomerular mesangial cell, Kidney proximal tubule brush border cell, Macula densa cell; reproductive system cells, including Spermatozoon, Sertoli cell, Leydig cell, Ovum; and other cells, including Adipocyte, Fibroblast, Tendon cell, Epidermal keratinocyte (differentiating epidermal cell), Epidermal basal cell (stem cell), Keratinocyte of fingernails and toenails, Nail bed basal cell (stem cell), Medullary hair shaft cell, Cortical hair shaft cell, Cuticular hair shaft cell, Cuticular hair root sheath cell, Hair root sheath cell of Huxley's layer, Hair root sheath cell of Henle's layer, External hair root sheath cell, Hair matrix cell (stem cell), Wet stratified barrier epithelial cells, Surface epithelial cell of stratified squamous epithelium of cornea, tongue, oral cavity, esophagus, anal canal, distal urethra and vagina, basal cell (stem cell) of epithelia of cornea, tongue, oral cavity, esophagus, anal canal, distal urethra and vagina, Urinary epithelium

cell (lining urinary bladder and urinary ducts), Exocrine secretory epithelial cells, Salivary gland mucous cell (polysaccharide-rich secretion), Salivary gland serous cell (glycoprotein enzyme-rich secretion), Von Ebner's gland cell in tongue (washes taste buds), Mammary gland cell (milk secretion), Lacrimal gland cell (tear secretion), Ceruminous gland cell in ear (wax secretion), Eccrine sweat gland dark cell (glycoprotein secretion), Eccrine sweat gland clear cell (small molecule secretion). Apocrine sweat gland cell (odoriferous secretion, sex-hormone sensitive), Gland of Moll cell in eyelid (specialized sweat gland), Sebaceous gland cell (lipid-rich sebum secretion), Bowman's gland cell in nose (washes olfactory epithelium), Brunner's gland cell in duodenum (enzymes and alkaline mucus), Seminal vesicle cell (secretes seminal fluid components, including fructose for swimming sperm), Prostate gland cell (secretes seminal fluid components), Bulbourethral gland cell (mucus secretion), Bartholin's gland cell (vaginal lubricant secretion), Gland of Littre cell (mucus secretion), Uterus endometrium cell (carbohydrate secretion), Isolated goblet cell of respiratory and digestive tracts (mucus secretion), Stomach lining mucous cell (mucus secretion), Gastric gland zymogenic cell (pepsinogen secretion), Gastric gland oxyntic cell (hydrochloric acid secretion), Pancreatic acinar cell (bicarbonate and digestive enzyme secretion), Paneth cell of small intestine (lysozyme secretion), Type II pneumocyte of lung (surfactant secretion), Clara cell of lung, Hormone secreting cells, Anterior pituitary cells, Somatotropes, Lactotropes, Thyrotropes, Gonadotropes, Corticotropes, Intermediate pituitary cell, Magnocellular neurosecretory cells, Gut and respiratory tract cells, Thyroid gland cells, thyroid epithelial cell, parafollicular cell, Parathyroid gland cells, Parathyroid chief cell, Oxyphil cell, Adrenal gland cells, chromaffin cells, Leydig cell of testes, Theca interna cell of ovarian follicle, Corpus luteum cell of ruptured ovarian follicle, Granulosa lutein cells, Theca lutein cells, Juxtaglomerular cell (renin secretion), Macula densa cell of kidney, Metabolism and storage cells, Barrier function cells (Lung, Gut, Exocrine Glands and Urogenital Tract), Kidney, Type I pneumocyte (lining air space of lung), Pancreatic duct cell (centroacinar cell), Nonstriated duct cell (of sweat gland, salivary gland, mammary gland, etc.), Duct cell (of seminal vesicle, prostate gland, etc.), Epithelial cells lining closed internal body cavities, Ciliated cells with propulsive function, Extracellular matrix secretion cells, Contractile cells; Skeletal muscle cells, stem cell, Heart muscle cells, Blood and immune system cells, Erythrocyte (red blood cell), Megakaryocyte (platelet precursor), Monocyte, Connective tissue macrophage (various types), Epidermal Langerhans cell, Osteoclast (in bone), Dendritic cell (in lymphoid tissues), Microglial cell (in central nervous system), Neutrophil granulocyte, Eosinophil granulocyte, Basophil granulocyte, Mast cell, Helper T cell, Suppressor T cell, Cytotoxic T cell, Natural Killer T cell, B cell, Natural killer cell, Reticulocyte, Stem cells

and committed progenitors for the blood and immune system (various types), Pluripotent stem cells, Totipotent stem cells, Induced pluripotent stem cells, adult stem cells, Sensory transducer cells, Autonomic neuron cells, Sense organ and peripheral neuron supporting cells, Central nervous system neurons and glial cells, Lens cells, Pigment cells, Melanocyte, Retinal pigmented epithelial cell, Germ cells, Oogonium/Oocyte, Spermatid, Spermatocyte, Spermatogonium cell (stem cell for spermatocyte), Spermatozoon, Nurse cells, Ovarian follicle cell, Sertoli cell (in testis), Thymus epithelial cell, Interstitial cells, and Interstitial kidney cells.

[0201] The cell (or target cell) can be engineered to comprise (or exhibit) any one of the systems or compositions as disclosed herein or can be treated by any one of the methods disclosed herein in vitro or ex vivo, then administered to the subject, e.g., to treat a condition of the subject. For example, any subject modified cell product can be administered to the subject to treat a condition of a bodily tissue of the subject. In some cases, the cell can be resident inside the subject's body, and any of the systems or compositions thereof can be administered to the subject, to contact the cell by the systems/compositions (e.g., to engineer the cell with the systems/compositions).

EXAMPLES

Example 1. Gene expression modulation

[0202] Gene expression can be modulated in a cell by utilizing a system or a method described herein. In some cases, the gene being modulated by the system or the method can be a mutant allele that can cause a disease or condition in a subject. In some cases, the gene being modulated can be a non-disease causing variant (e.g. a wild-type allele). In some embodiments, the gene expression can be modulated by the system or the method described herein by both decreasing the expression of the mutant allele in a cell and simultaneously increasing expression of the wild-type allele. In some cases, the wild-type allele is encoded by at least one of the heterologous polynucleotides described herein. **FIG. 1** illustrates an exemplary construct encoding the dCas and the actuator moiety (effector). dCas can be coupled with a transcription repressor for decreasing expression of the expression of the mutant allele of the endogenous target gene in the cell. **FIG. 2** illustrates a schematic for treating retinitis pigmentosa with the system described herein. AAV can be engineered to deliver an exemplary construct via subretinal or intravitreal injection to a subject in need thereof.

[0203] The modulation of the endogenous target gene expression by the system or method described herein can be used to treat a disease or condition in a subject. A subject suspected of having a disease or condition associated with mutation of the endogenous target gene can be first screened for the presence of a mutant allele of the endogenous target gene. Afterwards, the system described herein can be administered to the subject to simultaneously decrease expression

of the mutant allele of endogenous target gene and increase expression of the non-disease causing allele of endogenous target gene.

Example 2. PRPF31 expression modulation

[0204] PRPF31 expression can be modulated in a cell by utilizing a system or a method described herein. In some cases, the PRPF31 is a mutant allele of PRPF31. In some cases, the mutant PRPF31 can cause disease in a subject. In some cases, the PRPF31 is a non-disease causing variant of PRPF31 (e.g. wild type allele of PRPF31). In some embodiments, the PRPF31 expression can be modulated by the system or the method described herein by both decreasing the expression of the mutant allele of PRPF31 in a cell and simultaneously increasing expression of the wild type allele of PRPF31. In some cases, the wild type allele of PRPF31 is encoded by at least one of the heterologous polynucleotides described herein. **FIG. 3** illustrates exemplary target endogenous polynucleotide sequences (e.g., transcripts) that can be targeted by the gRNA of the system and the method described herein for decreasing or increasing the expression of PRPF31. The modulation of the PRPF31 expression by the system or method described herein can be used to treat a disease or condition in a subject. A subject suspected of having a disease or condition associated with PRPF31 mutation can be first screened for the presence of PRPF31 variant (e.g., a mutant allele of PRPF31).

Example 3. Expression cassettes (or constructs).

[0205] The systems as provided herein can be delivered to a cell via one or more expression cassettes encoding one or more components of the systems. The one or more expression cassettes can comprise a vector, such as a viral vector (e.g., an AAV vector comprising two Inverted terminal repeats (ITRs)). **FIGs. 4A-4F** schematically illustrate examples of a single vector encoding one or more components the system as provided herein.

[0206] **FIG. 4A** shows a schematic representation of a construct in which a RNA Pol II promoter drives the expression of a nuclear localization signal (NSL), a deactivated Cas (dCas), a protein linker (e.g., a GS linker), a modulator (e.g., a transcriptional effector), and Poly A signal placed at the 3'-end of the modulator. Additionally, the construct includes a RNA Pol III promoter that drives the expression of a guide RNA (gRNA) exhibiting specific binding against an endogenous target gene (e.g., encoding PRPF31)-terminator sequence, which is placed at the 3'-end of the poly A signal.

[0207] **FIG. 4B** shows a schematic representation of a construct in which a RNA Pol III promoter that drives the expression of a guide RNA (gRNA) exhibiting specific binding against

an endogenous target gene (e.g., encoding PRPF31)-terminator sequence, and a terminator sequence placed at the 3'-end of the gRNA. Additionally, the construct includes, downstream of the 3'-end of the terminator sequence, a RNA Pol II promoter drives the expression of a nuclear localization signal (NSL), a deactivated Cas (dCas), a protein linker (e.g., a GS linker), a modulator (e.g., a transcriptional effector), and Poly A signal placed at the 3'-end of the modulator.

[0208] FIG. 4C shows a schematic representation of a construct in which a RNA Pol II promoter drives the expression of a nuclear localization signal (NSL), a deactivated Cas (dCas), a protein linker (e.g., a GS linker), a modulator (e.g., a transcriptional effector), and Poly A signal placed at the 3'-end of the modulator. Additionally, the construct includes a RNA Pol III promoter that drives the expression of a guide RNA (gRNA) exhibiting specific binding against an endogenous target gene (e.g., encoding PRPF31)-terminator sequence. The RNA Pol III promoter (and the components under the control of such promoter) is downstream and on the opposite strand of the RNA Pol II promoter (and the components under the control of such promoter).

[0209] FIG. 4D shows a schematic representation of a construct similar to that shown in **FIG. 4A**. As compared to that in **FIG. 4A**, the RNA Pol II promoter (and the components under the control of such promoter) in the construct of **FIG. 4D** is further downstream away from the 5' ITR and towards the 3' ITR.

[0210] FIG. 4E shows a schematic representation of a construct similar to that shown in **FIG. 4B**. As compared to that in **FIG. 4B**, the RNA Pol II promoter (and the components under the control of such promoter) in the construct of **FIG. 4E** is further upstream towards the 5' ITR and away from the 3' ITR.

[0211] FIG. 4F shows a schematic representation of a construct similar to that shown in **FIG. 4C**. As compared to that in **FIG. 4C**, the RNA Pol II promoter (and the components under the control of such promoter) in the construct of **FIG. 4F** is further upstream towards the 5' ITR and away from the 3' ITR.

Example 4. gRNA screening.

[0212] Populations of retinal pigment epithelium (RPE) cells (e.g., human RPE cells) were engineered with a library of gRNAs against different portions of PRPF31 gene, along with a dCas-modulator (e.g., transcription activator such as VPR) fusion protein (e.g., via transduction with AAV vectors comprising of one of the constructs discussed in Example 3). Subsequently, expression level of PRPF31 (e.g., a non-disease causing PRPF31 variant) in the RPE cells was measured (e.g., via a gene expression assay such as enzyme-linked immunoassay (ELISA) or

reverse transcription-polymerase chain reaction (RT-PCR)) (**FIG. 5A**), to rank different gRNAs (e.g., with different spacer sequences and/or different viral vector construct designs) for the ability to induce the expression level of PRPF31. As shown in **FIG. 5B**, certain gRNAs (e.g., from a library of gRNAs derived from **Table 5**) yielded greater expression of PRPF31 in the RPE cells.

Example 5. In vitro functional assay.

[0213] Induced pluripotent stem cells (iPSCs) derived from cells of healthy or non-diseased subjects/patients (control iPSC) and iPSCs derived from cells of a RP11 subjects/patients (e.g., RP11-iPSCs) can be differentiated into RPE cells (control RPE cells and RP11-RPE cells, respectively) using an established differentiation protocol. PRPF31 co-localization with small nuclear ribonucleoproteins (snRNPs) and/or ADP-ribosylation factor-like protein 13B (ARL13B) can be assessed for the iPSC-derived RPEs, to confirm PRPF31 localization in both splicing complexes and cilia.

Example 6. Ciliogenesis.

[0214] To measure cilia length and incidence in RP11-RPE cells, immunostaining against the cilia marker ARL13B can be performed on one or more members of the following: (i) wild type RPE cells (e.g., control RPE cells or isolated RPE cells) that is not engineered with the system of the present disclosure (e.g., a complex comprising dCas-activator and a guide RNA against PRPF31), (ii) RP11-RPE cells (e.g., RP11-iPSCs or isolated mutant RPE cells) that are not engineered with the system of the present disclosure, and (iii) RP11-RPE cells that are engineered with the system of the present disclosure. Both cilia incidence and cilia length can be measured via imaging (e.g., ImageExpress Micro) for statistical analysis.

[0215] **FIGs. 6A** and **6B** show examples images of spliceosome normal RPE cells (**FIG. 6A**, with cilium indicated by four arrows) and RPE cells with defected spliceosome (**FIG. 6B**). The RPE cells with spliceosome defect display decreased cilium.

Example 7. Detection of splicing events.

[0216] PRPF31 mutations can lead to impaired pre-mRNA splicing of key components involved in the splicing process itself. To evaluate the restoration of splicing events after transducing the system of the present disclosure (e.g., a complex comprising dCas-activator and a guide RNA against PRPF31) to mutant RPE, gene expression assays (e.g., RT-PCR) can be performed in RPE cells and retinal organoids derived from either control iPSCs or RP11-iPSCs. Expression

level and change thereof of one or more key genes involved in cilia formation and/or outer segments of photoreceptors can be assessed accordingly. Non-limiting examples of such key genes can include RPGR, RPGRIP1L, CNOT3, intraflagellar transport (IFT122), actin filament organization, centrosome and focal adhesion (SORBS1), and pre-mRNA 3'-end processing (CPSF1).

Example 8. In vivo assay.

[0217] Gene delivery constructs (e.g., AAV constructs as shown in Example 3) in saline or a blank saline can be injected into the eye in vivo to treat or ameliorate an ocular disease (e.g., RP11).

[0218] For example, mice (e.g., age 4–5 weeks) can be anaesthetized and their pupils can be dilated. Mice can then be placed under an operating microscope. Upon visualizing the fundus of the mouse eye, a guide hole can be made through the sclera behind the iris (e.g., at an angle) towards the optic nerve. A micro-injection syringe can then be inserted into the hole and AAV (or blank saline) can be delivered to the vitreous body (for intravitreal injections), or the needle can be used to pierce the temporal retina and inject AAV to the subretinal space (for subretinal injections).

[0219] Effects of the AAV treatment can be assessed in later time points via visualization (e.g., funduscopy, optical coherence tomography (OCT)), retinal function assay (e.g., electroretinogram (ERG)), immunohistochemistry (IHC), gene expression analysis or sequencing of retinal cells extracted from the mice, etc.

Table 1. Non-limiting examples of a Cas protein or a derivative thereof.

SEQ ID NO	Amino acid sequence of Cas protein or a derivative thereof
1	MAKNTITKTLKLRIVRPYN SAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSVCYTRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKIGEK SAWMLNLSIDVPKIDKGVDPSSIIGGIDVGVSPLVCAINNAFSRYSISDNDLFHFNKKMFA RRRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPFKCEKCNFKENADYNAAALNISNPKLKSTKEEP

2	MEVQKTVMKTLRLRPLYSQEIEKEIKEEKERRKQAGGTGELDGGFYKKLEKKHSEM FSFDRLNLLLNLQREIAKVYNHAISELYATIAQGNKSNKHYYISSIVYNRAYGYFYNNAYI ALGICSKVEANFRSNELLTQQSALPTAKSDNFPIVLHKQKGAEGEDGGFRISTEGSDLIFEI PIPFYEYNGENRKEPYKWVKKGGQKPVLLKLSTFRRQRNKGWAKDEGTD AEIRKVTE GKYQVSQIEINRGKKLGEHQKWFANFSIEQPIYERKPNRSIVGGLDVGIRSLVCAINNSF SRYSDVSDNVFKFSKQVFAFRRLLSKNSLKRKGHGAHKL EPISTEMTEKNDKFRKKIIE RWAKEVTNFFVKNQVGIVQIEDLSTMKDREDHFFNQYLRGFWPYYQMOTLIENKLKEY GIEVKRVQAKYTSQLCSNPNCRYWNNYFNFEYRKVNKFPKFKCEKCNLEISADYNAAR NLSTPDIEKFVAKATKGINLPEK
3	MIKVYRYEIVKPLDLDWKEFGTILRQLQQETRFALNKATQLAWEWMGFSSDYKDNHGE YPKSKDILGYTNVHGYAYHTIKTKAYRLNSGNLSQTIKRATDRFKAYQKEILRGDMSIPS YKRDIPDLIKENISVNRNMNHGDYIASLSLLSNPAKQEMNVKRKISVIIIIRGAGKTIMDRI LSGEYQVSASQIIHDDRKNKWYLNISYDFEPQTRVLDLNKIMGIDLGVAVAVYMAFQHT PARYKLEGGEIENFRRQVESRRISMLRQGKYAGGARGGHGRDKRIKPIEQLRDKIANFRD TTNHRYSTRYIVDMAIKEGCGTIQMEDLTNIRDIGSRFLQNWTTYDLQKKIYKAE EAGIK VIKIDPQYTSQRCSECGNIDSGNRIGQAIFKCRACGYEANADYNAARNIAIPNIDKIIAESI K
4	MEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAYCTTQVERNACLFCKARKLD DKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEH YLSRVCYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKG GQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRK RNKGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVD PSIIGGIAVGVRSLVCAINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGA KNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQMENLESMKRKEDSYFNIRL RGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHF KCEKCNFKENAAAYNAALNISNPCLKSTKERP
5	MNNREKIALEKNKDKVKEACSKHLKVAAYCTTQVERNACLFCKARKLDDKFYQKLRG QFPDAVFWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRVCYRR AAELFKNAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISN HNSDFIIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNKGWSKDE GTEAEIKKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVG RSPLVCAINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILT EKSERFRKKLIERWACEIADFFIKNKVGT VQMENLESMKRKEDSYFNIRLRGFWPYAEM QNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHF KCEKCNFKE NAAYNAALNISNPCLKSTKERP
6	MAKN TITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE

	LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIKIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENA
7	MAKN TITKTLKL RIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIKIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFK
8	MAKN TITKTLKL RIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVNACLFCKARKLDD KFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHY LSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGG QYTGFEISNHNSDFIIKIPFGRWQVKKEIDKYRPWEKFDFEQVQKSPKPISLLLSTQRRKR NGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPS IIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFARRRILLKKNRHKRAGHGAKN KLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRG FWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFP HFKC EKC NFKENAAYNAALNISNPKLKSTKERP
9	MAKN TITKTLKL RIVRPYNSAEVEKIVADEKNNREKIALECKARKLDDKFYQKLRGQFP DAVFWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNS DFIIKIPFGRWQVKKEIDKYRPWEKFDFEQVQKSPKPISLLLSTQRRKRNGWSKDEGTE AEIKKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP L VCAINNAFSRYSISDNDLFHFNKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSE RFRKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKI EFKLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFP HFKCEKCNFKENAAY NAALNISNPKLKSTKERP
10	MAKN TITKTLKL RIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK

	SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFD EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFYRKKKNKFPFKCEKCNFKENAAYNAALNISNPKLKSTKERP
12	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRVCYRRAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFI KIPFGRWQVKKEIDKYRPWEKFD EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFYRKKKNKFPFKCEKCNFKENAAYNAALNISNPKLKSTKERP
13	MNMSKTTISVKLKIIDLSEKKEFLDNFYNEYAKATTCQLRIRRLRNTHWLKGKKEKSS KKWIFESGICDLGENKELVNEDRNSGEPAKICKRCYNGRYGNQMIRKLFVSTKKREVQ ENMDIRRVAKLNNTHYHRIPEEAFDMIKAAADTAEKRRKKNVEYDKKRQMEFIEMFNDE KKRAARPKPNERETRYVHISKLESPSKGYTLNGIKRKIDGMGKKIERAEKGLSRKKIFG YQGNRIKLDNWNVRFDLAESEITIPSLFKEMKLRLITGPTNVHSSKQIYFAEWFERINKQP NNYCYLIRKTSNGKYEYYLQYTYEAEVEANKEYAGCLGVDIGCSKLAAVYYDSKNK KAQKPIEFTNPIKKIKMRREKLIKLLSRVKVRHRRRKLMLQSKTEPIIDYTCHKTARKIVE MANTAKAFISMENLETGIKQKQARETKKQKFYRNMFLFRKLSKLIEYKALLKGIVY VKPDYTSQTCSSCGADKEKTERPSQAIFRCLNPTCRYYQRDINADFNAAVNIKKALNN TEVVTTLL
14	MPSETYITKTLCLKLIPSDDEKQALENYFITFQRAVNFAIDRIVDIRSSFRYLNKNEQFPAV CDCCGKKEKIMYVNISNKTFFKPSRNQKDRYTKDIYTIKPNHICKTCYSGVAGNMFIR KQMYPNDKEGWKVSRSYNIKVNAPGLTGTEYAMAIRKAISILRSFEKRRRNAERRIIEYE KSKKEYLELIDDEKGTNKIVVLEKEGHQVRVKRYKHKNWPEKWQGISLNKAKSKVK DIEKRIKKLKEWKHPTLNRPYVELHKNNVRIVGYETVELKLGKNMYTIHFASISNLRKPF RKQKKKSIEYLKHLTLALKRNLETYPSIIKRGKNFFLQYPVRVTVKVPKLTKNFKAFGI DRGVNRLAVGCIISKDGKLTNKNIFFFHGKEAWAKENRYKKIRDRLYAMAKKLRGDKT KKIRLYHEIRKKFRHKVKYFRRNYLHNISKQIVEIAKENTPTVIVLEDRLYLRETTYRGK GRSKKAKKTNYKLNTFTYRMLIDMIKYAAEEAGVPVMIIDPRNTSRKCSKCGYVDENN RKQASFKCLKCGYSLNADLNAAVNIKAFYECPTFRWEEKLHAYVCSEPK
15	MKSFKLKLPTDEQNVLLNEVFCKWASLCTRMASKGHDKERLAPPDSSGNYFNKTQLN QVNTDVTDHMGALLESASQKERA VEKVKRRLKLISDMLSEPNLRDVSQQKPTTFRPLE

	WVKEGLLKTKYHTVHYWQKECDKLTQKQKERMETIEKIKKGKITFKPTKMSLHQNCFS LSFGKGTFSMRPFSDTKRGINLDMLTAPIQPAIGKNDGKSSLSKEFIARNIENYIIFSIHSQ LFGLSRSEELLLNAKKEELVAKRDAMLKKKSDSLSKKIKELEKIVGRKITDSERSEIMSQ GGKLSSEKFSEDNSYLKTLKVLAKDIIGREELFRLKKYPIVIRKPLNERKKLKNLKPDEW EYYLQLSYDELEKKEFTPKTIMGIDRGLKHILAIAIYDPVQNKVKNMLIPNPILGWKWK LRKIKRSIQHMERRIRAQQNAHV PENQLKKRLKSIENKIDYYYHNVSQRILNLAHDFKSA IVVEDLQNMKQHGRKKSKGLRGLNYALS NF DY GKIMGLVKYKAESENVPLLT VLPAGT SQNCAYCLLYGKEQGN YVRNNVNSKIGKCKLHGEIDADINAARTIAICYHKNINEPKPY GERKTFKRK
16	MKYTKVMRYQIIKPLNAEWDELGMVLRDIQKETRAALNKTIQLCWEYQGFSADYKQIH GQYPKPKDVLGYTSMHGYAYDRLKNEFSKIASNLSQTIKRAVDKWNSDLKEILRGDRS IPNFRKDCPIDIVKQSTKIQKCNDGYVLSLGLINREYKNELGRKNGVFDVLIKANDKTQQ TILERIINGDYTYTASQIINHKNKWFINTYQFETKETALDPNNVMGVDLGIVYPVYIAFN NSLHRYHIKGGIEFRFRRQVEKRKRELLNQGKYCGDGRKGHGYATRKSIESISDKIARF RDTCNHKYSRFIVDMALKHNCGIIQMEDLTGISKESTFLKNWTTYDLQKIEYKAREAGI QVIKIEPQYTSQRCSKCGYIDKENRQEATFKCIECGFKTNADYNAARNIAIPNIDKIIRKT LKMQ
17	MTLLVKVVKIHLISEQFDKAGNRIDYEEVNKILWELQKQTREAKNKT VQLLWEWNNFS SDYVKASGIYPKAKDIFGYSSVHGQANKELRTKLALNSSNLSTTTMDVCKNFNTYKKEV WKGKRSVPSYKSDQPLDLHKDSIKLIYENNEFYVRLALLKKAFAKYGFKDGRFKMQ VKDNSTKTILERC FDEVYKINASKLLYDQKKKKWKLNLSY SFDNKNISELDKEKILGVD VGVNCPLVASVFGDRDRFIKGGIEKFRKSVEARRRSMLEQTKYCGDGRIGHGRKKRT EPALNIGDKIARFRD TTNHKYSRALIEYAVKKGCGTIQMEKLTGITSKSDRFLKDWTYYD LQTKIENKAKEVGINVVYIAPKYTSQRCSKCGYIHKDNRPNQAKFRCLECDFESNADYN ASQNIKIDKIIKDLQKQSEVQVNENK
18	MGESVKAIKLKILDMFLDPECTKQDDNWRKDLSTMSRFCAEAGNMCLRDLYNYFSMP KEDRISSKDLYNAMYHKTKLLHPELPGKVANQIVNHAKDVWKRNAKLIYRNQISMPTY KITTAIRLQNNIYKLIKNNKYIIDVQLYSKEYSKDSGKGTHRYFLVAVRDSSTRMIFDR IMSKDHIDSSKSYTQGQLQIKKDHQGWYCIIPYTFPTHETVLDPKVMGVDLGVAKAV YWAFNSSYKRGCIDGGEIEHFRKMIRARRVSIQNQIKHSGDARKGHGRKRALKPIETLSE KEKNFRDTINHRYANRIVEAAIKQCGGTI QIENLEGIADTTGSKFLKNWPYYDLQTKIVN KAKEHGITVVA INPQYTSQRCSMCGYIEKTNRSSQAVFECKQCGYGSRTICINCRHVQVS GDVCEECCGIVKKENVNADYNAAKNISTPYIDQIIMEKCLELGIPYRSITCKECGHIQASG NTCEVCGSTNILKPKKIRKAK
19	MITVRKIKLTIMGDKDTRNSQYKWIRDEQYNQYRALNMGMTYLA VNDILYMNESGLEI RTIKDLKDCEKDIDKNKKEIEKLTARLEKEQNKNSSSEKLDEIKYKISLVENKIEDYKLIK IVELNKILEETQKERMDIQKEFKKYVDDLYQVLDKIPFKHLDNKS LVTQRIKADIKSDK

	SNGLLK GERSIR NYKRN FPLMTRGRDLKFYDDND D IEIKWMEGIKFKVILGNRIKNSLE LRHTLHKVIEGKYKICDSSLQFDKNNNLILNLTL DIPIDIVNKKVSGRVVGV DLGLKIPAY CALNDVEYIKKSIGRIDDFLKVRTQM QSRRRLQIAIQSAKGGKGRVNKLQALERFAEK EKNFAKTYNHFLSSNIVKFAVSNQAEQINMELLSLKETQNK SILRNWSYYQLQTMIEYK AQREGIKVKYIDPYHTSQTCSKCGNYEEGQRESQADFICKKCGYKVNADYNAARNIAM SNKYITKKEESKYKIKESMV
20	MAKN TITKTLKL RIVRPYNSAEVEKIVADEKERRKQAGGTGELDGGFYKKLEKKHSEMF SFDRLNLLLNLQLQKQAAEIYNQSLIELYYE IFIKGGK GIANASSVEHYLSRV CYRRAAE LFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGF EISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKR NKGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFP HFKCEKCNFKENAAYN AALNISNPCLKSTKERP
21	MEVQKTVMKTL SLRILRPLYSQEIEKEIKEEKERRKQAGGTGELDDKFYQKLRGQFPDA VFWQEISEIFRQLQKQAAEIYNQSLIELYYE IFIKGGK GIANASSVEHYLSRV CYRRAAE LFK KNAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGF EISNHNSD FIIKIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKR NKGWSKDEGTEA EIKKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLV CAINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSER FRKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIE FKLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFP HFKCEKCNFKENAAYN AALNISNPCLKSTKERP
22	MAKN TITKTLKL RIVRPLYSQEIEKEIKEEKERRKQAGGTGELDDKFYQKLRGQFPDAVF WQEISEIFRQLQKQAAEIYNQSLIELYYE IFIKGGK GIANASSVEHYLSRV CYRRAAE LFKN AAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGF EISNHNSDFIIK IPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKR NKGWSKDEGTEAEIK KVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVCA INNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERFR KKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFK LKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFP HFKCEKCNFKENAAYN ALNISNPCLKSTKERP
23	MIKVYRYEIVKPLDL DWQEISEIFRQLQKQAAEIYNQSLIELYYE IFIKGGK GIANASSVEHY LSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGG QYTGF EISNHNSDFIIKIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKR NKGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPS

	IIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKN KLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRG FWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHFKE EKC NF KENAAYNAALNISNP KLKSTKERP
24	MITVRKIKL TIMGDKDTRNSQYK WIRDEQYNQYRALNMGMTYLA VND AVFWQEISEIF RQLQKQAAEIYNQSLIELYYE I FIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLR SKI KSNFRLKELKNMKSGLP TTKSDNFPIPLVKQKGGQYTGFEISNHNSDFI IKIPFGRWQ VKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGD YQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSR YSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERW ACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEI RKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHFKEKC NF KENAAYNAALNISNP K LKSTKERP
25	MGESVKA I K L K I L D M F L D P E C T K Q D D N W Q E I S E I F R Q L Q K Q A A E I Y N Q S L I E L Y Y E I F I K G K G I A N A S S V E H Y L S R V C Y R R A A E L F K N A A I A S G L R S K I K S N F R L K E L K N M K S G L P T T K S D N F P I P L V K Q K G G Q Y T G F E I S N H N S D F I I K I P F G R W Q V K K E I D K Y R P W E K F D F E Q V Q K S P K P I S L L L S T Q R R K R N K G W S K D E G T E A E I K K V M N G D Y Q T S Y I E V K R G S K I C E K S A W M L N L S I D V P K I D K G V D P S I I G G I A V G V R S P L V C A I N N A F S R Y S I S D N D L F H F N K K M F A R R R I L L K K N R H K R A G H G A K N K L K P I T I L T E K S E R F R K K L I E R W A C E I A D F F I K N K V G T V Q M E N L E S M K R K E D S Y F N I R L R G F W P Y A E M Q N K I E F K L K Q Y G I E I R K V A P N N T S K T C S K C G H L N N Y F N F E Y R K K N K F P H F K C E K C N F K E N A A Y N A A L N I S N P K L K S T K E R P
26	MKYTKVMRYQIIKPLNAEWQEISEIFRQLQKQAAEIYNQSLIELYYE I F I K G K G I A N A S S V E H Y L S R V C Y R R A A E L F K N A A I A S G L R S K I K S N F R L K E L K N M K S G L P T T K S D N F P I P L V K Q K G G Q Y T G F E I S N H N S D F I I K I P F G R W Q V K K E I D K Y R P W E K F D F E Q V Q K S P K P I S L L L S T Q R R K R N K G W S K D E G T E A E I K K V M N G D Y Q T S Y I E V K R G S K I C E K S A W M L N L S I D V P K I D K G V D P S I I G G I A V G V R S P L V C A I N N A F S R Y S I S D N D L F H F N K K M F A R R R I L L K K N R H K R A G H G A K N K L K P I T I L T E K S E R F R K K L I E R W A C E I A D F F I K N K V G T V Q M E N L E S M K R K E D S Y F N I R L R G F W P Y A E M Q N K I E F K L K Q Y G I E I R K V A P N N T S K T C S K C G H L N N Y F N F E Y R K K N K F P H F K C E K C N F K E N A A Y N A A L N I S N P K L K S T K E R P
27	MTLLVKVVKIHLISEQFDKAGNRIDYEEISEIFRQLQKQAAEIYNQSLIELYYE I F I K G K G I A N A S S V E H Y L S R V C Y R R A A E L F K N A A I A S G L R S K I K S N F R L K E L K N M K S G L P T T K S D N F P I P L V K Q K G G Q Y T G F E I S N H N S D F I I K I P F G R W Q V K K E I D K Y R P W E K F D F E Q V Q K S P K P I S L L L S T Q R R K R N K G W S K D E G T E A E I K K V M N G D Y Q T S Y I E V K R G S K I C E K S A W M L N L S I D V P K I D K G V D P S I I G G I A V G V R S P L V C A I N N A F S R Y S I S D N D L F H F N K K M F A R R R I L L K K N R H K R A G H G A K N K L K P I T I L T E K S E R F R K K L I E R W A C E I A D F F I K N K V G T V Q M E N L E S M K R K E D S Y F N I R L R G F W P Y A E M Q N K I E F K L K Q Y G I E I R K V A P N N T S K T C S K C G H L N N Y F N F E Y R K K N K F P H F K C E K C N F K E N A A Y N A A L N I S N P K L K S T K E R P

28	MAKNTITKTLKLRIVRPYYSQEIEKIVAEKKNRREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPCLKSTKERP
29	MAKNTITKTLKLRIVRPYYSAEVEKIVAEKKNRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKLRGQFPDAVFWQEISEIFRQLQKQAREIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPCLKSTKERP
30	MAKNTITKTLKLRIVRPYYSAEIEKIVADEKNRREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRKQFPDAVFWQEISEIFRQLQKQAREIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPCLKSTKERP
31	MAKNTITKTLKLRIVRPYNSQEVEKIVAEKKNRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRKQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPCLKSTKERP

32	MAKNTITKTLKL RIVRPYNSQEVEKIVAE EKNNREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKKLRKQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSE RFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIE FKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAY NAALNISNPCLKSTKERP
33	MAKNTITKTLKL RIVRPYNSQEVEKIVAE EKNNREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKKLRGQFPDAVFWQEISEIFRQLQKQAREIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAE IKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSE RFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIE FKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAY NAALNISNPCLKSTKERP
34	MAKNTITKTLKL RIVRPYYSAEVEKIVAE EKNNREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKKLRKQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAE IKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSE RFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIE FKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAY NAALNISNPCLKSTKERP
35	MAKNTITKTLKL RIVRPYNSAEIEKIVADE KNRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKKLRKQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAE IKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSE RFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIE FKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAY NAALNISNPCLKSTKERP

36	MAKNTITKTLKL RIVRPYNSAEIEKIVADEKNRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKKLRGQFPDAVFWQEISEIFRQLQKQAREIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPCLKSTKERP
37	MAKNTITKTLKL RIVRPYNSQEIEKIVAEENRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPCLKSTKERP
38	MAKNTITKTLKL RIVRPYNSAEIEKIVAEENRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPCLKSTKERP
39	MAKNTITKTLKL RIVRPYNSAEVEKIVAEENRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKKLRKQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPCLKSTKERP

40	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYKKLRKQFPDAVFWQEISEIFRQLQKQAREIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPFKCEKCNFKENAAYNAALNISNPCLKSTKERP
41	MAKNTITKTLKLRIVRPYYS AEIEKIVADEKNRREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKL RKQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPFKCEKCNFKENAAYNAALNISNPCLKSTKERP
42	MAKNTITKTLKLRIVRPYYS AEIEKIVAEENRREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKL RKQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPFKCEKCNFKENAAYNAALNISNPCLKSTKERP
43	MAKNTITKTLKLRIVRPYYS AEIEKIVAEENNREKIALDKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKL RGQFPDAVFWQEISEIFRQLQKQAREIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPFKCEKCNFKENAAYNAALNISNPCLKSTKERP

44	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAARLFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFKISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP
45	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAALFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFKISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP
46	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAGLFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFKISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP
47	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAARLFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFRISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP

48	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAALFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFRISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP
49	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAGLFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFRISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP
50	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAARLFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFSISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP
51	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAALFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFSISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGTQVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP

52	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAGLFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFSISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF RQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIRKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPKLKSTKERP
53	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLENFNKMF A RRRILLKKNRHKRGGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPKLKSTKERP
54	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIEGGDLFHFNKMF A RRRILLKKNRHKRAGHGAKNKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPKLKSTKERP
55	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIEGGDLENFNKMF A RRRILLKKNRHKRGGHGGRD KKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPKLKSTKERP

56	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIEGGDLENFNKKMFA RRRILLKKNRHKRAGHGAKNKLK PITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAA YNAALNISNPKLKSTKERP
57	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKKMFAR RRRILLKKNRHKRAGHGRDKKLKPIEQLTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAA YNAALNISNPKLKSTKERP
58	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLENFNKKMFA RRRILLKKNRHKRAGHGRDKKLK PITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAA YNAALNISNPKLKSTKERP
59	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLEHFNKKMFA RRRILLKKNRHKRKGHGAKNKLKPIETLTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAA YNAALNISNPKLKSTKERP

60	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIDGGDLEHFNKKMFA RRRILLKKNRHKRKGHGAKNKLK PITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPKLKSTKERP
61	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIDGGDLFHFNKKMFA RRRILLKKNRHKRAGHGAKNKLK PIETLTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPKLKSTKERP
62	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIDGGDLEHFNKKMFA RRRILLKKNRHKRKGHGAKNKLK PIETLTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPKLKSTKERP
63	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNSFSRYSIDSNDLFKFNKKMFAR RRILLKKNRHKRKGHGAKNKLK PITELTEKSERFRKKLIERWACEIADFFIKNKVGT VQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFP HFKCEKCNFKENAAAYNAALNISNPKLKSTKERP

64	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNSFSRYSIDSNDLFHFNKKMFAR RRILLKKNRHKRAGHGAAHKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAYNAALNISNPCLKSTKERP
65	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNSFSRYSIDSNDLFHFNKKMFAR RRILLKKNRHKRAGHGAAHKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAYNAALNISNPCLKSTKERP
66	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSIDSNDLFHFNKKMFAR RRILLKKNRHKRAGHGAAHKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAYNAALNISNPCLKSTKERP
67	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNSFSRYSIDSNDLFHFNKKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAYNAALNISNPCLKSTKERP

68	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNSFSRYSISDNDLFKFNKKMFAR RRILLKKNRHKRKGHGAKNKLKPITELTEKSERFRKKLIERWACEIADFFIKNKVGTVMQ ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYNAALNISNPCLKSTKERP
69	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIDSNDLFKFNKKMFAR RRILLKKNRHKRKGHGAKNKLKPITELTEKSERFRKKLIERWACEIADFFIKNKVGTVMQ ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYNAALNISNPCLKSTKERP
70	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFKFNKKMFAR RRILLKKNRHKRKGHGAHKLKPITELTEKSERFRKKLIERWACEIADFFIKNKVGTVMQ ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYNAALNISNPCLKSTKERP
71	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIKGGDLERFNKKMFA RRRILLKKNRHKRKGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVMQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYNAALNISNPCLKSTKERP

72	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIKGGDLERFNKKMFA RRRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAA YNAALNISNPKLKSTKERP
73	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIKGGDLEKFNKKMFA RRRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAA YNAALNISNPKLKSTKERP
74	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIKGGDLFHFNKKMFA RRRILLKKNRHKRAGHGGRKKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAA YNAALNISNPKLKSTKERP
75	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLEKFNKKMFA RRRILLKKNRHKRAGHGGRKKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFP HFKCEKCNFKENAA YNAALNISNPKLKSTKERP

76	MAKNTITKTLKL RIVRPYN SAEVEKIV ADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSIKGGDLEKFNKKMFA RRRILLKKNRHKRAGHGRRKKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCG HLNNYFNFEYRKKNKFPFKCEKCNFKENAAAYNAALNISNPKLKSTKERP
77	MAKNTITKTLKL RIVRPYN SAEVEKIV ADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWAKEIADFFIKNKVGT VQM EDLSTMKRKEDSYFNIRLRGFWPYEMQNKIEFKLKQYGIEIRKVAPNNTSQLCSKCGH LNNYFNFEYRKKNKFPFKCEKCNFKENAAAYNAARNISTPDIKSTKERP
78	MAKNTITKTLKL RIVRPYN SAEVEKIV ADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGT VQM EDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQLCSKCGH LNNYFNFEYRKKNKFPFKCEKCNFKENAAAYNAALNISNPDIKSTKERP
79	MAKNTITKTLKL RIVRPYN SAEVEKIV ADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWAKEIADFFIKNKVGT VQM EDLSTMKRKEDSYFNIRLRGFWPYEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPFKCEKCNFKENAAAYNAALNISNPKLKSTKERP

80	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQLCSKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKENAAYNAALNISNPDIKSTKERP
81	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQLCSKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKENAAYNAALNISTPDIKSTKERP
82	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM EDLSTMKRKEDSYFNIRLRGFWPYEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKENAAYNAALNISNPCLKSTKERP
83	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQM EDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQLCSKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKENAAYNAALNISTPDIKSTKERP

84	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKENAAAYNAARNISTPDIKSTKERP
85	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWSRYIADFFIKNKVGTVM EDLESMKRKEDSYFNIRLRGFWPYEMQNKIEFKLKQYGIRKVAPNNTSQRC SKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKENAAAYNAARNISNPNIKSTKERP
86	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACYIADFFIKNKVGTVM EDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQRC SKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKENAAAYNAARNISNPNIKSTKERP
87	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACYIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQRC SKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKRNAAYNAARNISNPKLKSTKERP

88	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACYIADFFIKNKVGTVM EDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYNAARNISNPNIKSTKERP
89	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWARYIADFFIKNKVGTVM EDLESMKRKEDSYFNIRLRGFWPYEMQNKIEFKLKQYGIKIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYNAALNISNPCLKSTKERP
90	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQRC SKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKRNAAYNAARNISNPNIKSTKERP
91	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM EDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQRC SKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYNAARNISNPNIKSTKERP

92	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKRNAAYNAARNISNPNIKSTKERP
93	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWANRIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIKIRKVAPNNTSQRC SKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKRNAAYNAAKNISNPCLKSTKERP
94	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQRC SKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKRNAAYNAAKNISNPCLKSTKERP
95	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIKIRKVAPNNTSQRC SKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKRNAAYNAAKNISNPCLKSTKERP

96	MAKNTITKTLKL RIVRPYN SAEVEKIV ADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWANRIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIRKVAPNNTSQRCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAYNAALNISNPKLKSTKERP
97	MAKNTITKTLKL RIVRPYN SAEVEKIV ADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWANRIADFFIKNKVGTVM ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQRCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKRNAAYNAALNISNPKLKSTKERP
98	MAKNTITKTLKL RIVRPYN SAEVEKIV ADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWSRFIADFFIKNKVGTVM EDLESMKRKEDSYFNIRLRGFWPYEMQNKIEFKLKQYGIEIRKVAPNNTSQRCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAYNAARNISNPNIKSTKERP
99	MAKNTITKTLKL RIVRPYN SAEVEKIV ADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIDVQLYSKEYSKDSGKGTHRYFLLS TQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKID KGVDPSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKMFARRILLKKNRHKRA GHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVMENLESMKRKEDSY FNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKN KFPHPKCEKCNFKENAAYNAALNISNPKLKSTKERP

100	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIASLSLLSNPAKQEMNVKRKISLLS TQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKID KGVDPSSIIGGIAVGVRSP LVCAINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRA GHGAKNLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSY FNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKN KFPHFKEKCNFKENAAAYNAALNISNPKLKSTKERP
101	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVP IYERKPNRSIVGGLAVGIRSPLVCAINNSFSRYSVDSNDVFKFSKQVFAF RRRLLSKNSLKRKGHGA AHKLEPITEMTEKNDKFRKKI IERWAKEVTNFFVKNQVGIVQ IEDLSTMKDREDHFFNQYLRGFWPYYQMQLTIENKLKEYGIEVKRVQAKYTSQLCSNPN CRYWNNYFNFEYRKVNKFPKFKCEKCNLEISAAAYNAARNLSTPDIEKFVAKATKGINLP EK
102	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVP THETVLDPDKVMGVALGVAKAVYWAFNSSYKRGCIDGGEIEHFRKMI RARRVSIQNQIKHSGDARKGHGRKRALKPIETLSEKEKNFRDTINHRYANRIVEAAIKQG CGTIQIENLEGIADTTGSKFLKNWPYYDLQTKIVNKAKEHGITVVAINPQYTSQRCSMCG YIEKTNRSSQAVFECKQCGYGSRTICINCRHVQVSGDVCEECCGIVKKENVNAAYNAAK NISTPYIDQIIMEKCLELGIPYSITCKEKGHIQASGNTCEVCGSTNILKPKK
103	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPQTRVLDLNKIMGIALGVAVAVYMAFQHTPARYKLEGGEIENFRRQVE SRRISMLRQGKYAGGARGGHGRDKRIKPIEQLRDKIANFRDTTNHRSRYIVDMAIKEG

	CGTIQMEDLTNIRDIGSRFLQNWYYDLQQKIHYKAEAGIKVIKIDPQYTSQRCSECGNI DSGNRIGQAIFKCRACGYEANAAYNAARNIAIPNIDKIIAESIK
104	MAKN'TITKTLKLRIVRPYN'SAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNRKGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPIDIVNKKVSGRVVGVALGLKIPAYCALNDVEYIKKSIGRIDDFLKVRTQ MQSRRRRLQIAIQSAKGGKGRVNLQALERFAEKEKNFAKTYNHFLSSNIVKFAVSNQA EQINMELLSLKETQNK SILRNWSYYQLQTMIEYKAQREGIKVKYIDPYHTSQTCSKCGN YEEGQRESQADFICKKCGYKVNAAYNAARNIAMS NKYITKKEESKYKIKESMV
105	MAKN'TITKTLKLRIVRPYN'SAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNRKGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVETKETALDPNNVMGVALGIVYPVYIAFNNSLHRYHIKGGEIERFRRQVE KRKRELLNQGKYCGDGRKGHGYATRTKSIESISDKIARFRDTCNHKYSRFIVDMALKHN CGIIQMEDLTGISKESTFLKNWYYDLQQKIEYKAREAGIQVIKIEPQYTSQRC SKCGYID KENRQE QATFKCIECGFKTNAAYNAARNIAIPNIDKIIRKTLKMQ
106	MAKN'TITKTLKLRIVRPYN'SAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNRKGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVNRSIVGGLAVGIRSP LVC AINNSFSRYSVDSNDVFKFSKQVFA FRRRLSKNSLKRKGHGA AHKLEPITEMTEKNDKFRKKIIERWAKEVTNFFVKNQVGIV QIEDLSTMKDREDHFFNQYLRGFWPYYQMQLTIENKLKEYGIEVKRVQAKYTSQLCSNP NCRYWNNYFNFEYRKVNKFPKFKCEKCNLEISAAYNAARNLSTPDIEKFVAKATKGINL PEK
107	MAKN'TITKTLKLRIVRPYN'SAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAE LFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNRKGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPDKVMGVALGVAKAVYWAFNSSYKRGCIDGGEIEHFRKMI RARRVSIQNQIKHSGDARKGHGRKRALKPIETLSEKEKNFRDTINHRYANRIVEAAIKQG

	CGTIQIENLEGIADTTGSKFLKNWPYYDLQTKIVNKAKEHGITVVAINPQYTSQRCSMCG YIEKTNRSSQAVFECKQCGYGSRTICINCRHVQVSGDVCEECCGIVKKENVNAAYNAAK NISTPYIDQIIMEKCLELGIPYRSITCKEKGHIQASGNTCEVCGSTNILKPKK
108	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDLNKIMGIALGVAVAVYMAFQHTHPARYKLEGGEIENFRRQV ESRRISMLRQGYAGGARGGHGRDKRIKPIEQLRDKIANFRDTTNHRYSRIVDMAIKE GCGTIQMEDLTNIRDIGSRFLQNWTYYDLQKQIYKAAEEAGIKVIKIDPQYTSQRCSECGN IDSGNRIGQAIFKCRACGYEANAAYNAARNIAIPNIDKIIAESIK
109	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPNVMGVALGIVYPVYIAFNNSLHRYHIKGGIEIERFRRQVE KRRRELLNQGYCGDGRKGHGYATRKTIESISDKIARFRDTCNHKYSRFIVDMALKHN CGIIQMEDLTGISKESTFLKNWTYYDLQKQIEYKAREAGIQVIKIEPQYTSQRCSKCGYID KENRQEATFKCIECGFKTNAAYNAARNIAIPNIDKIIRKTLKMQ
110	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDKEKILGVAVGVCNPLVASVFGDRDRFIKGGIEIEKFRKSVEA RRRSMLEQTKYCGDGRIGHGRKKRTEPALNIGDKIARFRDTTNHKYSRALIEYAVKKGC GTIQMEKLTGITSKSDRFLKDWTYDDLQTKIENKAKEVGINVVYIAPKYTSQRCSKCGYI HKDNRPNQAKFRCLECFESNAAYNASQNIKIDKIIKDLQKQESEVQVNENK
111	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVCAINNAFSRYSISDNDLFHFNKKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVM

	ENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSKTCSKCGH LNNYFNFEYRKKNKFPKFKCEKCNFKENAAYNAALNI
112	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
113	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
114	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISTPDIKSTKERP
115	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQRCSCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AARNISNPNIKSTKERP
116	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKFNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPCLKSTKERP
117	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKFNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNI
118	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKFNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
119	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKFNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
120	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLFFKNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AARNI
121	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLFFKNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
122	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFFKNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
123	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFFKNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
124	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRVCYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISTPDIKSTKERP
125	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRVCYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AARNISNPNIKSTKERP
126	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRVCYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPCLKSTKERP
127	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRVCYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
128	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
129	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
130	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AARNI
131	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNI
132	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPDIKSTKERP
133	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
134	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISTPDIKSTKERP
135	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQRCSCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AARNISNPNIKSTKERP
136	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPCLKSTKERP
137	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQ MEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNI
138	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
139	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NK KM FARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQ MEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
140	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AARNI
141	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNSFSRYSISDNDL FKF NKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
142	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNAFSRYSISDNDL FHF NKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
143	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNAFSRYSISDNDL FHF NKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
144	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISTPDIKSTKERP
145	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFP HFKCEKCNFKENAAYN AARNISNPNIKSTKERP
146	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQ MENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFP HFKCEKCNFKENAAYN AALNISNPCLKSTKERP
147	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVR SPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
148	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
149	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
150	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AARNI
151	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHKCEKCNFKENAAYN AALNI
152	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHKCEKCNFKENAAYN AALNISNPDIKSTKERP
153	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
154	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISTPDIKSTKERP
155	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQRCSCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AARNISNPNIKSTKERP
156	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPPLVC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPPLKSTKERP
157	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPPLVC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNI
158	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPPLVC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
159	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPPLVC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
160	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPPLVC AINNSFSRYSISDNDLFFKNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AARNI
161	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPPLVC AINNSFSRYSISDNDLFFKNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
162	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPPLVC AINNAFSRYSISDNDLFFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
163	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPPLVC AINNAFSRYSISDNDLFFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
164	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISTPDIKSTKERP
165	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AARNISNPNIKSTKERP
166	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPCLKSTKERP
167	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
168	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAGVRSPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
169	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAGVRSPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
170	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAGVRSPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AARNI
171	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAGVRSPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNI
172	MAKNTITKTLKLRIVRPYNSQEIEKIVAAEEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLKFKNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPDIKSTKERP
173	MAKNTITKTLKLRIVRPYNSQEIEKIVAAEEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLKFKNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
174	MAKNTITKTLKLRIVRPYNSQEIEKIVAAEEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLKFKNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISTPDIKSTKERP
175	MAKNTITKTLKLRIVRPYNSQEIEKIVAAEEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLKFKNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQRCCKGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AARNISNPNIKSTKERP
176	MAKN'TITKTLKLRIVRPYNSQEIEKIVAEERKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNSFSRYSISDNDL FKFNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSCKGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPCLKSTKERP
177	MAKN'TITKTLKLRIVRPYNSQEIEKIVAEERKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNSFSRYSISDNDL FKFNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSCKGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNI
178	MAKN'TITKTLKLRIVRPYNSQEIEKIVAEERKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNSFSRYSISDNDL FKFNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSCKGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
179	MAKN'TITKTLKLRIVRPYNSQEIEKIVAEERKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFD FEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNSFSRYSISDNDL FKFNKKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
180	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AARNI
181	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPVLC AINNSFSRYSISDNDLFFKNKMFARRRILLKKNRHKRKGHGAKNKLKPITELTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNI
182	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFFKNKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
183	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFFKNKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
184	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISTPDIKSTKERP
185	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AARNISNPNIKSTKERP
186	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPCLKSTKERP
187	MAKNTITKTLKLRIVRPYNSQEIEKIVAEERKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
188	MAKNTITKTLKLRIVRPYNSQEIEKIVAAEEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
189	MAKNTITKTLKLRIVRPYNSQEIEKIVAAEEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AALNI
190	MAKNTITKTLKLRIVRPYNSQEIEKIVAAEEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQRC SKCGHLNNYFNFEYRKKNKFPFKCEKCNFKENAAYN AARNI
191	MAKNTITKTLKLRIVRPYNSQEIEKIVAAEEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNI
192	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNSFSRYSISDNDLFKNKKMFAR RRILLKKNRHKRKGHGAKNKLKPITELTEKSERFRKKLIERWACEIADFFIKNKVGTVMQ EDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQLCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPDIKSTKERP
193	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKNNREKIALEKNKDKVKEACSKHLKVAAY CTTQVERNACLFCKARKLDDKFYQKLRGQFPDAVFWQEISEIFRQLQKQAAEIYNQSLIE LYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFKNAAIASGLRSKIKSNFRLKELKNMK SGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFIIPFGRWQVKKEIDKYRPWEKFDF EQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEIKKVMNGDYQTSYIEVKRGSKICEKS AWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFAR RRILLKKNRHKRAGHGAKNKLKPITILTEKSERFRKKLIERWACEIADFFIKNKVGTVMQ EDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEFKLKQYGIEIRKVAPNNTSQLCSKCGH LNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPDIKSTKERP
194	MAKNTITKTLKLRIVRPYNSAEIEKIVADEKERRKQAGGTGELDDKFYKKLRKQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLP TTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDFEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVMQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAYN AALNISNPDIKSTKERP
195	MAKNTITKTLKLRIVRPYNSAEIEKIVADEKERRKQAGGTGELDDKFYKKLRKQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAALFK NAAIASGLRSKIKSNFRLKELKNMKSGLP TTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDFEQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSP LVC AINNAFSRYSISDNDLFHFNKKMFARRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVMQENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
196	MAKNTITKTLKLRIVRPYNSAEIEKIVADEKERRKQAGGTGELDDKFYKKLRKQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
197	MAKNTITKTLKLRIVRPYNSAEIEKIVADEKERRKQAGGTGELDDKFYKKLRKQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPDIKSTKERP
198	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMEDLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPKFKCEKCNFKENAAYN AALNISNPCLKSTKERP
199	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRV CYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSIIGGIAVGVRSPVLC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF

	KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYN AALNISNPCLKSTKERP
200	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAGVRSPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSKTCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYN AALNISNPDIKSTKERP
201	MAKNTITKTLKLRIVRPYNSAEVEKIVADEKERRKQAGGTGELDDKFYQKLRGQFPDAV FWQEISEIFRQLQKQAAEIYNQSLIELYYEIFIKGKGIANASSVEHYLSRVCYRRAAELFK NAAIASGLRSKIKSNFRLKELKNMKSGLPTTKSDNFPIPLVKQKGGQYTGFEISNHNSDFII KIPFGRWQVKKEIDKYRPWEKFDQVQKSPKPISLLLSTQRRKRNGWSKDEGTEAEI KKVMNGDYQTSYIEVKRGSKICEKSAWMLNLSIDVPKIDKGVDPSSIIGGIAGVRSPLVC AINNAFSRYSISDNDLFHFNKKMFARRRILLKKNRHKRAGHGAKNKLKPITILTEKSERF RKKLIERWACEIADFFIKNKVGTVQMENLESMKRKEDSYFNIRLRGFWPYAEMQNKIEF KLKQYGIEIRKVAPNNTSQLCSKCGHLNNYFNFEYRKKNKFPHPKCEKCNFKENAAAYN AALNISNPDIKSTKERP

Table 2. Non-limiting examples of a guide nucleic acid scaffold sequence or a derivative thereof.

SEQ ID NO	Guide nucleic acid (NA) scaffold sequence (without spacer)
500	GGCTTCACTGATAAAGTGGAGAACCGCTTACCAAAAAGCTGTCCCTTAGGGGATTAG AACTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCG GAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATGCAAC
501	GAACCGCTTACCAAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCT GCTTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAA ATTCATTTGAATGAAGGAATGCAAC
502	AACCGCTTACCAAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTG CTTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAA TTCATTTGAATGAAGGAATGCAAC

503	ACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATGCAAC
504	CCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTTCATTTGAATGAAGGAATGCAAC
505	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAGCAATAAGGAATGCAAC
506	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAGAAAGGAATGCAAC
507	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATCTTCGGATTAAGGAATGCAAC
508	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTGCAAAGGAATGCAAC
509	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCGTTAAGGAATGCAAC
510	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATGCAAAGGAATGCAAC
511	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAGCAATTAAGGAATGCAAC
512	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAGAACTTGAGTGAAGGTGGGCTGCTGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAGAAAGGAATGCAAC
513	GGCTTCACTGATAAAGTGGAGAACCGCTTCACTTAGAGTGAAGGTGGGCTGCTTGCA TCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATT TGAATGAAGGAATGCAAC
514	GGCTTCACTGATAAAGTGGAGAACCGCTTCACTTCGAGTGAAGGTGGGCTGCTTGCA TCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATT TGAATGAAGGAATGCAAC

515	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCTTCGGTGAAGGTGGGCTGCTTGCAT CAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTT GAATGAAGGAATGCAAC
516	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCTTAGGAGTGAAGGTGGGCTGCTTG CATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCA TTTGAATGAAGGAATGCAAC
517	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCTTCGGAGTGAAGGTGGGCTGCTTG CATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCA TTTGAATGAAGGAATGCAAC
518	GGCTTCACTGATAAAGTGGAGAACCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCT TGCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATT CATTTGAATGAAGGAATGCAAC
519	ACCGCTTCACCAAAAGCTGTCTTAGGGATTAGAACTTGAGTGAAGGTGGGCTGCTT GCATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTC ATTTGAATGAAGGAATGCAAC
520	ACCGCTTCACCAAAAGCTGTCTTAGGATTAGAACTTGAGTGAAGGTGGGCTGCTTGC ATCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCAT TTGAATGAAGGAATGCAAC
521	ACCGCTTCACCAAAAGCTGTTTAGATTAGAACTTGAGTGAAGGTGGGCTGCTTGCAT CAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTT GAATGAAGGAATGCAAC
522	ACCGCTTCACCAAAAGCTGTTAGTTAGAACTTGAGTGAAGGTGGGCTGCTTGCATCA GCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGA ATGAAGGAATGCAAC
523	ACCGCTTCACCAAAAGCTTTAGAGAACTTGAGTGAAGGTGGGCTGCTTGCATCAGCC TAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGAATG AAGGAATGCAAC
524	ACCGCTTCACCAAAAGCTTCGGCACTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTA ATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGAATGAA GGAATGCAAC
525	ACCGCTTCACCAAAAGTTCGCACTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAAT GTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGG AATGCAAC
526	ACCGCTTCACCAAAATTCGTCTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGT CGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAA TGCAAC

527	ACCGCTTCACCAAGTTCGCTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTGCG AGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATG CAAC
528	ACCGCTTCACCAATTCGTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTGCGAG AAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATGCA AC
529	ACCGCTTCACCATTCGTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTGCGAGAA GTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATGCAAC
530	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAG AACTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTGCGAGAAGTGCTTTCTTCG GAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATGCAAC
-	TT
532	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAG AACTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTGCGAGAAGTGCTTTCTTCG GAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATGCAAC
-	TTTTA
534	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAG AACTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTGCGAGAAGTGCTTTCTTCG GAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATGCAAC
-	TTTTG
536	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAG AACTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTGCGAGAAGTGCTTTCTTCG GAAAGTAACCCTCGAAACAAATTCATTTGAATGAAGGAATGCAAC
537	TTTTATTTTT
538	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCGAGTGAAGGTGGGCTGCTTGCATC AGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATTTG AATGAAGGAATGCAAC
539	TTTTATTTTT
540	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCAAAAGCTGTCCCTTAGGGGATTAG AACTTGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTGCGAGAAGTGCTTTCTTCG GAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
541	GGCTTCACTGATAAAGTGGAGAACCGCTTCACTTAGAGTGAAGGTGGGCTGCTTGCA TCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAATTCATT TGAATGAAGGAATGCAAC
542	TTTTATTTTT

543	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCGAGTGAAGGTGGGCTGCTTGCATC AGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCCTCGAAACAAATTCATTTG AATGAAGGAATGCAAC
544	TTTTATTTTT
545	GGCTTCACTGATAAAGTGGAGAACCGCTTCACTTAGAGTGAAGGTGGGCTGCTTGCATC TCAGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAG GAATGCAAC
546	GGCTTCACTGATAAAGTGGAGAACCGCTTCACCGAGTGAAGGTGGGCTGCTTGCATC AGCCTAATGTCGAGAAGTGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGA ATGCAAC
547	ACCGCTTCACTTAGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGTGC TTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
548	ACCGCTTCACCGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGTGCTT TCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
549	ACCGCTTCACTTAGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGTGC TTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
550	TTTTATTTTT
551	ACCGCTTCACCGAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGTGCTT TCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
552	TTTTATTTTT
553	ACCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAA GTGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
554	ACCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAA GTGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
-	TTTTA
555	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
556	CCGCTTCACGCTTAGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
557	CCGCTTCACTCTTAGGAAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
558	CCGCTTCACGTTTAGACAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC
559	GCCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAA GTGCTTTCTTCGGAAAGTAACCCCTCGAAACAAAGAAAGGAATGCAAC

560	GGCCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGA AGTGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
561	CGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGT GCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
562	GCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGTG CTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
563	GGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGT GCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
564	GCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGTG CTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
550	TTTTATTTTT
565	GGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAGT GCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
550	TTTTATTTTT
566	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAGGAATGCAAC
567	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAGGAATGCAAC
568	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGGGAATGCAAC
569	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGGAATGCAAC
570	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAGGAATGCAAC
571	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAGGAATGCAAC
572	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACGGAATGCAAC
573	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAAGGAATGCAAC
574	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAGGAATGCAAC
575	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAGGAATGCAAC
576	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAGAATGCAAC

577	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGGAATGCAAC
578	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAATGCAAC
579	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGAATGCAAC
580	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAGAATGCAAC
581	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAATGCAAC
582	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAGAATGCAAC
583	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAATGCAAC
584	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
585	CCGCTTCACGCTTCGGCAGTGAAGGTAGGCTGCTTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
586	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCCAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
587	GGCTTCACGCTTCGGCAGTGAAGGTAGGCTGCTTGCATCAGCCTAATGTCGAGAAGT GCTTTCTTCGGAAAGTAACCCTCGAAACAAGGAATGCAAC
588	GGCTTCACGCTTCGGCAGTGAAGGTGGGCTGCTTGCATCAGCCCAATGTCGAGAAGT GCTTTCTTCGGAAAGTAACCCTCGAAACAAGGAATGCAAC
589	CCGCTTCACTCTTAGGAAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAAAGGAATGCAAC
590	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGGAATGCAAC
591	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGGAATGCAAC
592	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAGGAATGCAAC
593	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAGAATGCAAC
594	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGGAATGCAAC

595	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAAAGAATGCAAC
596	CCGCTTCACGCTTCGGCAGTGAAGGTGGGCTGATTGCATCAGCCTAATGTCGAGAAG TGCTTTCTTCGGAAAGTAACCCTCGAAACAGAATGCAAC

Table 3. Exemplary list of guide RNA scaffold fragment sequences.

SEQ ID NO	Guide nucleic acid (NA) scaffold fragment sequence
597	CCGCTTCACGCTTCGGCAGTGAAGGTGGGC
598	CCGCTTCACTCTTAGGAAGTGAAGGTGGGC
599	GAAAGTAACCCTCGAAACAAAGAATGCAAC
600	AAGTAACCCTCGAAACAAAGGGAATGCAAC
601	AAAGTAACCCTCGAAACAAAGGAATGCAAC

Table 4. Exemplary gRNA spacer nucleic acid sequence

SEQ ID NO	Target gRNA spacer sequence (without scaffold)	Target genomic location
700	TCACTCTCGCGCCGTTATAGAGG	chr19:54115722
701	TGTCTCGACGCCCCGTCGTCCGG	chr19:54115533
702	AATCGCTGTGGCCGGACGACGGG	chr19:54115544
703	TTAGCAGGATCGGTCCACAGCGG	chr19:54115573
704	GCGCGAGAGTGAGACGTCATCGG	chr19:54115733
705	GCGGCCAGGAACATCGCGAGCGG	chr19:54115676
706	GCCGAAGCGAAGCCGGAAACAGG	chr19:54115416
707	GAGAGAGGCGGTAAGCCGCGAGG	chr19:54115614
708	CGGCCAGGAACATCGCGAGCGGG	chr19:54115677
709	CGGTAAGCCGCGAGGAGGAAAGG	chr19:54115606
710	GTCCCCGCTCGCGATGTTCTCTGG	chr19:54115680
711	GGCCAGGAACATCGCGAGCGGGG	chr19:54115678
712	GATTCTCTGCTTAGCAGGATCGG	chr19:54115563
713	TAGCAGGATCGGTCCACAGCGGG	chr19:54115574
714	ACAGCGATTCTCTGCTTAGCAGG	chr19:54115558
715	TCCTGTTTCCGGCTTCGCTTCGG	chr19:54115415
716	TTGTTGTTTCAAATGCGGCCAGG	chr19:54115662
717	TATTAGCACCTGGCACTAGGCGG	chr19:54115636

718	GAATCGCTGTGGCCGGACGACGG	chr19:54115545
719	GCTTCCGGAGGTTCGCGGAGATGG	chr19:54115266
720	GAGTAGCTTTGCCTCTATAACGG	chr19:54115711
721	GGCCTTTAAGCAGGGATTCTGGGG	chr19:54115454
722	GCTTAACCTACAGATTGAAGAGG	chr19:54115310
723	TGCTAAGCAGAGAATCGCTGTGG	chr19:54115556
724	CGGCGTTAGACAAGAAAGCAAGG	chr19:54115475
725	AAGGGACTCACGTCCCCTGTGG	chr19:54115587
726	GGGACAAGTAGTTTTGCGAACGG	chr19:54115332
727	ATCGCTGTGGCCGGACGACGGGG	chr19:54115543
728	GGGCCACGTCCCGCCGTCTCGG	chr19:54115369
729	GACGTGGCCCTAGGCCTTGTGGG	chr19:54115382
730	AGAGGCGGTAAAGCCGCGAGGAGG	chr19:54115611
731	GGACGTGGCCCTAGGCCTTGTGG	chr19:54115381
732	GCAAACTACTTGTCCCATCAGG	chr19:54115338
733	GTCGGAAGCTCTGAGGCCCGGGG	chr19:54115288
734	AATTCCATCTCCGCGACCTCCGG	chr19:54115262
735	TCCGCGACCTCCGGAAGCCCCGG	chr19:54115271
736	CGAGACAAGAAGAAAGACGTTGG	chr19:54115517
737	CAGGGATTCTGGGGTGGACGTGGG	chr19:54115444
738	AGGCAAAGCTACTCTCTGATTGG	chr19:54115702
739	GGGGTGGGCCGAAGCGAAGCCGG	chr19:54115423
740	GGTAAGCCGCGAGGAGGAAAGGG	chr19:54115605
741	ACAACAACCTTTATTAGCACCTGG	chr19:54115646
742	TGAAGAGGTCTGGAAGCTCTGAGG	chr19:54115295
743	ACCTGGCACTAGGCGGAGAGAGG	chr19:54115629
744	CTTTATTAGCACCTGGCACTAGG	chr19:54115639
745	CCGCGACCTCCGGAAGCCCCGGG	chr19:54115272
746	CAGCCGAGGACGGCGGGACGTGG	chr19:54115366
747	GGAGGTCTGCGGAGATGGAATTGG	chr19:54115260
748	CACCCCGAATCCCTGCTTAAAGG	chr19:54115452
749	GCAACTCAGAGGACTGGTTGCGG	chr19:54115495
750	AAGCAGAGAATCGCTGTGGCCGG	chr19:54115552
751	AGGGATTCTGGGGTGGACGTGGGG	chr19:54115443
752	ACGTTGGCAACTCAGAGGACTGG	chr19:54115501
753	TACAACTCCCACAAGGCCTAGGG	chr19:54115389

754	AGGCCTTTAAGCAGGGATTCGGG	chr19:54115455
755	GAAAGACGTTGGCAACTCAGAGG	chr19:54115506
756	CTTTAAGCAGGGATTCGGGGTGG	chr19:54115451
757	GCTCAGCAGCCGAGGACGGCGGG	chr19:54115360
758	GGACGGCGGGACGTGGCCCTAGG	chr19:54115373
759	TTCCGACCTCTTCAATCTGTAGG	chr19:54115304
760	GGCTCAGCAGCCGAGGACGGCGG	chr19:54115359
761	GGTCGGAAGCTCTGAGGCCCGGG	chr19:54115289
762	CCCATCAGGCTCAGCAGCCGAGG	chr19:54115352
763	CTGAGGCCCGGGGCTTCCGGAGG	chr19:54115278
764	GCAGGGATTCGGGGTGGACGTGG	chr19:54115445
765	TGGCACTAGGCGGAGAGAGGCGG	chr19:54115626
766	GTGAGTCCCTTTCCTCCTCGCGG	chr19:54115599
767	GCCTCTCTCCGCCTAGTGCCAGG	chr19:54115628
768	TCAGGCTCAGCAGCCGAGGACGG	chr19:54115356
769	AGAAAGCAAGGCCTTTAAGCAGG	chr19:54115463
770	GAAAGCAAGGCCTTTAAGCAGGG	chr19:54115462
771	ATTCGGGGTGGACGTGGGGGTGG	chr19:54115439
772	AGGAAACTACAACCTCCACAAGG	chr19:54115396
773	AACCTACAGATTGAAGAGGTCGG	chr19:54115306
774	GAGTTGTAGTTTCCTGTTTCCGG	chr19:54115404
775	AGGTCGGAAGCTCTGAGGCCCGG	chr19:54115290
776	CTACAACCTCCACAAGGCCTAGG	chr19:54115390
777	TCCTCGGCTGCTGAGCCTGATGG	chr19:54115353
778	TAAAGTTGTTGTTTCAAATGCGG	chr19:54115657
779	GGGATTCGGGGTGGACGTGGGGG	chr19:54115442
780	GCTCTGAGGCCCGGGGCTTCCGG	chr19:54115281
781	AAGGCCTTTAAGCAGGGATTCGG	chr19:54115456
782	TTCGGGGTGGACGTGGGGGTGGG	chr19:54115438
783	ATGGAATTGGAGCAGAGAGAAGG	chr19:54115247
784	TGGAATTGGAGCAGAGAGAAGGG	chr19:54115246
785	GGAGCAGAGAGAAGGGTATGTGG	chr19:54115239

Table 5. Exemplary gRNA spacer nucleic acid sequence

SEQ ID NO	Guide nucleic acid spacer sequence (without scaffold)	Strand	Chromosome	Start	End
786	TGTATTTTAGTAGAGATGGG	+	chr19	54113865	54113884
787	GTATTTTAGTAGAGATGGGG	+	chr19	54113866	54113885
788	TATTTTAGTAGAGATGGGGT	+	chr19	54113867	54113886
789	ATTTTAGTAGAGATGGGGTT	+	chr19	54113868	54113887
790	TAGAGATGGGGTTTCACCAT	+	chr19	54113875	54113894
791	TCTCCAAGGACCCAGGGAGT	+	chr19	54114975	54114994
792	CTCCAAGGACCCAGGGAGTC	+	chr19	54114976	54114995
793	GGGATTCGCACTTTTTCACA	+	chr19	54115210	54115229
794	ACAAGGGCTCAGCCACATAC	+	chr19	54115227	54115246
795	CAAGGGCTCAGCCACATACC	+	chr19	54115228	54115247
796	TTTTTTTGTTGTTTGTTTGT	+	chr19	54115898	54115917
797	TTTTTTGTTGTTTGTTTGTT	+	chr19	54115899	54115918
798	TGTTGTTTGTTTGTTTGAG	+	chr19	54115904	54115923
799	GTTGTTTGTTTGTTTGAGA	+	chr19	54115905	54115924
800	TTGTTTGTTTGTTTGAGAC	+	chr19	54115906	54115925
801	TGTTTGTTTGTTTGAGACG	+	chr19	54115907	54115926
802	GACGGAGTCTCGTTCTGTTG	+	chr19	54115923	54115942
803	TATTTTAGTAGAGACGGGG	+	chr19	54116076	54116095
804	ATTTTAGTAGAGACGGGGT	+	chr19	54116077	54116096
805	GTAGAGACGGGGTTTCCCCA	+	chr19	54116084	54116103
806	TAGAGACGGGGTTTCCCAT	+	chr19	54116085	54116104
807	CTGAGACGGAGTTTCGCTCT	+	chr19	54116220	54116239
808	GTATTTTAGTAGAGACGGG	+	chr19	54116378	54116397
809	TATTTTAGTAGAGACGGGG	+	chr19	54116379	54116398
810	ATTTTAGTAGAGACGGGGT	+	chr19	54116380	54116399
811	GTAGAGACGGGGTTTCACTA	+	chr19	54116387	54116406
812	TAGAGACGGGGTTTCACTAT	+	chr19	54116388	54116407
813	AGCCTGTTTACCCAAAAGTC	+	chr19	54116512	54116531
814	AGCTGGGTTTCCATACCTAG	+	chr19	54117522	54117541
815	AGAGATGGGGTTTCACCATG	+	chr19	54113876	54113895
816	CCAGGGGACCCCAAAACCGG	+	chr19	54114403	54114422
817	TCCCGCTGGCGCCGCCGCCG	+	chr19	54114497	54114516

818	GCTTGGTTTTTTGTTTTTTTG	+	chr19	54115886	54115905
819	AGAGACGGGGTTTCCCATG	+	chr19	54116086	54116105
820	AGAGACGGGGTTTCACTATG	+	chr19	54116389	54116408
821	GCCTGTTTACCCAAAAGTCT	+	chr19	54116513	54116532
822	CAAAAGTCTTAATATGCGCC	+	chr19	54116524	54116543
823	AAATAATTACTTCATTAATC	+	chr19	54116636	54116655
824	AATTAGCCGGGCGTGCTGGT	+	chr19	54116945	54116964
825	TACAAGGGACTCCTGTACAA	+	chr19	54117147	54117166
826	GATCAAATGCTCGCAGATGA	+	chr19	54117267	54117286
827	AGACCATTTCAAGCTTTAGG	+	chr19	54117313	54117332
828	CTGGGGAGAGGGGCTTAAAA	+	chr19	54117333	54117352
829	TTGTCAGGGAAGCTCCACGG	+	chr19	54114233	54114252
830	ATTTAATTCCCGCTGGCGCC	+	chr19	54114490	54114509
831	CCTGGCGGCGCTGAGAACTC	+	chr19	54114656	54114675
832	AGATCCAGCAGTCCAAACCC	+	chr19	54115053	54115072
833	ATTCGCACTTTTTCACAAGG	+	chr19	54115213	54115232
834	CCAATTCCATCTCCGCGACC	+	chr19	54115260	54115279
835	GGTTAAGCCGTTTCGCAAAAC	+	chr19	54115325	54115344
836	TTGCCAACGTCTTTCTTCTT	+	chr19	54115514	54115533
837	TAGCAGGATCGGTCCACAGC	+	chr19	54115574	54115593
838	AGAAGAAGTGTGTGAGGAAA	+	chr19	54115850	54115869
839	GCCCAGGTTGGAGAGCAGTG	+	chr19	54115942	54115961
840	TCAGCCTCCAGAGTAGCTGG	+	chr19	54116018	54116037
841	ACGGAGTTTCGCTCTTGTTG	+	chr19	54116225	54116244
842	ACTCAGAGGCAGAGGAAGTG	+	chr19	54117218	54117237
843	GTTTTGCAGCTGGGTTTCCA	+	chr19	54117515	54117534
844	GTAAAGCAGTTGCATGAACT	+	chr19	54115123	54115142
845	GCCGCAACCAGTCCTCTGAG	+	chr19	54115494	54115513
846	ACGGCGCGAGAGTGAGACGT	+	chr19	54115730	54115749
847	GTGGGGCGTGTGGGTCAGGG	+	chr19	54116736	54116755
848	GAAAAATTTAAAAATTAGCC	+	chr19	54116933	54116952
849	AAAAAAAAAAAAAAAAAATTA	+	chr19	54117685	54117704
850	TTCACCATGTTGGCCAGGCT	+	chr19	54113887	54113906
851	ATGTGAAGAGCTTGGGCTTT	+	chr19	54114152	54114171
852	TTCACCTGTGAGCGGTGCCA	+	chr19	54114170	54114189
853	TGTGGAAAGGCAGAGATTGC	+	chr19	54114365	54114384

854	TCGCACTTTTTTCAAGGGC	+	chr19	54115215	54115234
855	TTGTAGTTTCCTGTTTCCGG	+	chr19	54115407	54115426
856	CCACGGGGAGCGGGAGGCTG	+	chr19	54115814	54115833
857	CCTGGGTCTGGGAGAAGAAG	+	chr19	54115838	54115857
858	TGGGAGAAGAAGTGTGTGAG	+	chr19	54115846	54115865
859	AAGAAGTGTGTGAGGAAAAA	+	chr19	54115852	54115871
860	TACAGGCGCCCGCCACCACC	+	chr19	54116041	54116060
861	TACAGGCGCGCACCACCACG	+	chr19	54116343	54116362
862	TCCCAAAGTGCTGGGATTAC	+	chr19	54116462	54116481
863	TACAGGCATGAGCCACCGCT	+	chr19	54116479	54116498
864	TGTGGAAGGAGCAAAAATTG	+	chr19	54116566	54116585
865	GTGTGGGTGAGGGAGGTGGA	+	chr19	54116743	54116762
866	AGGGAGGTGGAACGTGAGAG	+	chr19	54116752	54116771
867	ACTAAGAAGTGAGGAGAGCC	+	chr19	54116797	54116816
868	GCCCAGGCGAAAGGATCGCT	+	chr19	54116858	54116877
869	GTGGTGAAAGTAAAGCAGCT	+	chr19	54117242	54117261
870	GAGGGGCTTAAAAAGGGGAA	+	chr19	54117340	54117359
871	GCATACAGGAGTGGTGCTGG	+	chr19	54117373	54117392
872	CAAGGTATGTGTGTCTTGCT	+	chr19	54117396	54117415
873	GGTGTCAAGTCAACAATGGC	+	chr19	54117458	54117477
874	TCCATACCTAGTTTGTTC	+	chr19	54117531	54117550
875	GCCAAGGTGGGTCGCTCACT	+	chr19	54117610	54117629
876	GCTCACTTGAGGTCAAGAGT	+	chr19	54117623	54117642
877	TGGTGGCAGGTGCCTGTAGT	+	chr19	54117713	54117732
878	TTGCAATGGCAGCCATAGGG	+	chr19	54114047	54114066
879	AGAGGAGAAGGGTCCAGAGA	+	chr19	54114074	54114093
880	GGCAGAGATTGCGGGGGGCG	+	chr19	54114373	54114392
881	CTGGACCCCCAGCCCCTTCT	+	chr19	54114736	54114755
882	CCCAGCCCCTTCTCCCACTG	+	chr19	54114743	54114762
883	GCAGTGGCGCGATCTCGGCT	+	chr19	54115956	54115975
884	GCAGTGGCGCGATCTCGGTT	+	chr19	54116258	54116277
885	GAGCAAAAATTGTTTCGCTAC	+	chr19	54116574	54116593
886	GTGAGAGCTGAAGGCTGAGG	+	chr19	54116765	54116784
887	CGGGAAGAGACCGATAATGT	+	chr19	54117178	54117197
888	ATGGGCTGGTGTCAAGTCAA	+	chr19	54117451	54117470
889	TTTGCAGCTGGGTTTCCATA	+	chr19	54117517	54117536

890	GTTTAGGCCAGGGGACCCCA	-	chr19	54114396	54114415
891	CAGGATGACCCCAGTCCATA	-	chr19	54115090	54115109
892	CCCCATCAGACCTCAGCGTA	-	chr19	54115148	54115167
893	TGTAGGTAAAGCCGTTTCGCA	-	chr19	54115321	54115340
894	GGAGAAGAAGTGTGTGAGGA	-	chr19	54115848	54115867
895	GAGAAGAAGTGTGTGAGGAA	-	chr19	54115849	54115868
896	TTTACAGCCTGTTTACCCA	-	chr19	54116506	54116525
897	CTGGGGATGTGGAAGGAGCA	-	chr19	54116559	54116578
898	TGGGGATGTGGAAGGAGCAA	-	chr19	54116560	54116579
899	TGAGACCCTGTCTCTACAGA	-	chr19	54116915	54116934
900	GAGACCCTGTCTCTACAGAA	-	chr19	54116916	54116935
901	TGTCTCTACAGAAAAATTTA	-	chr19	54116923	54116942
902	GTCTCTACAGAAAAATTTAA	-	chr19	54116924	54116943
903	CAGAGCAAGACATGGTCTCA	-	chr19	54117076	54117095
904	AGAGCAAGACATGGTCTCAA	-	chr19	54117077	54117096
905	GAGCAAGACATGGTCTCAAA	-	chr19	54117078	54117097
906	AGCAAGACATGGTCTCAAAA	-	chr19	54117079	54117098
907	GCAAGACATGGTCTCAAAAA	-	chr19	54117080	54117099
908	ACATGGTCTCAAAAAAAGA	-	chr19	54117085	54117104
909	GTCTCAAAAAAAGAAAAGA	-	chr19	54117090	54117109
910	AAAAAAAGAAAAGAAAAGA	-	chr19	54117095	54117114
911	AAAAAAAGAAAAGAAAAGAA	-	chr19	54117096	54117115
912	AAAAAAGAAAAGAAAAGAAA	-	chr19	54117097	54117116
913	AGGCTGGGGAGAGGGGCTTA	-	chr19	54117330	54117349
914	GGCTGGGGAGAGGGGCTTAA	-	chr19	54117331	54117350
915	AGTGAAACCTTGACTCTACA	-	chr19	54117665	54117684
916	GTGAAACCTTGACTCTACAA	-	chr19	54117666	54117685
917	TGAAACCTTGACTCTACAAA	-	chr19	54117667	54117686
918	GAAACCTTGACTCTACAAAA	-	chr19	54117668	54117687
919	AAACCTTGACTCTACAAAAA	-	chr19	54117669	54117688
920	AACCTTGACTCTACAAAAAA	-	chr19	54117670	54117689
921	ACCTTGACTCTACAAAAAAA	-	chr19	54117671	54117690
922	CCTTGACTCTACAAAAAAA	-	chr19	54117672	54117691
923	CTTGACTCTACAAAAAAA	-	chr19	54117673	54117692
924	TTGACTCTACAAAAAAA	-	chr19	54117674	54117693
925	TGACTCTACAAAAAAA	-	chr19	54117675	54117694

926	GACTCTACAAAAAAAAAAAAA	-	chr19	54117676	54117695
927	ACTCTACAAAAAAAAAAAAA	-	chr19	54117677	54117696
928	CTCTACAAAAAAAAAAAAA	-	chr19	54117678	54117697
929	TCTACAAAAAAAAAAAAA	-	chr19	54117679	54117698
930	CCAGACACATAATGCATGCT	-	chr19	54114284	54114303
931	AATGACTGAATATATAAGCT	-	chr19	54114305	54114324
932	TCAGGATGACCCCAGTCCAT	-	chr19	54115089	54115108
933	CCATAAAAGGGTTCTAAGGT	-	chr19	54115105	54115124
934	CCCCCATCAGACCTCAGCGT	-	chr19	54115147	54115166
935	TCCACCCCGAATCCCTGCTT	-	chr19	54115450	54115469
936	GCCTAGTGCCAGGTGCTAAT	-	chr19	54115638	54115657
937	CTGTCTCTACAGAAAATTT	-	chr19	54116922	54116941
938	AAGTGGGGGGTGGTGAAAGT	-	chr19	54117233	54117252
939	TAGGCTGGGGAGAGGGGCTT	-	chr19	54117329	54117348
940	CTATTGCAGTCATTCAGCCC	-	chr19	54114014	54114033
941	GGAGGAGAGGAGAAGGGTCC	-	chr19	54114069	54114088
942	AGAGGCCCCCGGGAGACATC	-	chr19	54114197	54114216
943	CAGGGGCCCGGCACAGGGCC	-	chr19	54114266	54114285
944	GGCTGGGCATGTGGAAAGGC	-	chr19	54114356	54114375
945	CCCGGAGCCCTGAGCCGCCC	-	chr19	54114549	54114568
946	TCCCACTGAACCAGGAGCCC	-	chr19	54114755	54114774
947	CCCCCAGCCCCTCCTCCCTC	-	chr19	54114814	54114833
948	GCTCCCAGCCCCTCCTCCCTC	-	chr19	54114849	54114868
949	AGCCCTCCTCCCTCAGACCC	-	chr19	54114855	54114874
950	CTCCCAGGATGCAAGAGTCC	-	chr19	54114937	54114956
951	ATGCAAGAGTCCAGACCTCC	-	chr19	54114945	54114964
952	CCAAGCCCCAACCCTCAACC	-	chr19	54114995	54115014
953	TTCGAACCTCCTAGTCTGTC	-	chr19	54115033	54115052
954	CATGAACTACAACCCCCATC	-	chr19	54115135	54115154
955	CTCCGGAAGCCCCGGGCCTC	-	chr19	54115279	54115298
956	ATCGCGAGCGGGGACCAATC	-	chr19	54115688	54115707
957	ATTCTTCTGCCTCAGCCTCC	-	chr19	54116007	54116026
958	GGTCTCCAACCTCCTGACCTC	-	chr19	54116420	54116439
959	AGCTTTGGTTGATGCCTGGC	-	chr19	54116602	54116621
960	GTGCCTTACTGTAGACACGC	-	chr19	54116669	54116688
961	ACTAAGAAGTGAGGAGAGCC	-	chr19	54116797	54116816

962	CAGTGAGACCCTGTCTCTAC	-	chr19	54116912	54116931
963	TGCACTGCAGCCTAGGTGAC	-	chr19	54117057	54117076
964	AACTCAAGTTTCTGCTACTC	-	chr19	54117202	54117221
965	AGTTTCTGCTACTCAGAGGC	-	chr19	54117208	54117227
966	TTTACTGATCAAATGCTCGC	-	chr19	54117261	54117280
967	TAATTTTTTTGTATTTTAGT	-	chr19	54113856	54113875
968	CCCCTCCTCCCTGAGATCCT	-	chr19	54114784	54114803
969	CCCCTCCTCCCTCAGACCGT	-	chr19	54114821	54114840
970	ATCGGTGAGCGACTAACGCT	-	chr19	54115751	54115770
971	CTAATTTTTGTATTTTAGT	-	chr19	54116066	54116085
972	TAGTTTTTTGTATTTTAGT	-	chr19	54116369	54116388
973	AAATTGTTTCGCTACCCTCTT	-	chr19	54116580	54116599
974	TGTGTGAAGTGCCTTACTGT	-	chr19	54116661	54116680
975	AAAGCAACCTCCTGTAGTCT	-	chr19	54116713	54116732
976	ATGGCCAAGCTAATGTCCTT	-	chr19	54117290	54117309
977	GTCAACAATGGCCACGTTGT	-	chr19	54117466	54117485
978	GATCCGCCTGCCTCAGCCTC	-	chr19	54113931	54113950
979	ACAAGCGTGAGCCACCATGC	-	chr19	54113965	54113984
980	CGTGAGCCACCATGCCCAGC	-	chr19	54113970	54113989
981	GGCTATTGCAGTCATTCAGC	-	chr19	54114012	54114031
982	CCCGGGAGACATCAGAGAAC	-	chr19	54114204	54114223
983	GTAGTTTAGGCCAGGGGACC	-	chr19	54114393	54114412
984	CTCCCGGAGCCCTGAGCCGC	-	chr19	54114547	54114566
985	CGCTGAGAACTCCTCAAAGC	-	chr19	54114664	54114683
986	GGGCCCTGGATCCTGGACCC	-	chr19	54114724	54114743
987	TGGACCCCCAGCCCCCTTCTC	-	chr19	54114737	54114756
988	TCTCCCACTGAACCAGGAGC	-	chr19	54114753	54114772
989	CTGAACCAGGAGCCCAGACC	-	chr19	54114760	54114779
990	AGATCCTAGAATCCAGGCCC	-	chr19	54114797	54114816
991	CAGACCGTAGAATCCAGCTC	-	chr19	54114833	54114852
992	CCAGCCCTCCTCCCTCAGAC	-	chr19	54114853	54114872
993	CAGAAGTCCAAGTCCGCAAC	-	chr19	54114874	54114893
994	CAACCCACCCTTCGCAGCAC	-	chr19	54114890	54114909
995	CAAGCCCTGACCCCCTCCTC	-	chr19	54114920	54114939
996	GACTTTTTCTCTCCAAGGAC	-	chr19	54114966	54114985
997	GGACCCAGGGAGTCCAAGCC	-	chr19	54114982	54115001

998	TCTCCTCCCTCAGGATGACC	-	chr19	54115080	54115099
999	AGTTGCATGAACTACAACCC	-	chr19	54115130	54115149
1000	CGTTCGCAAACTACTTGTC	-	chr19	54115333	54115352
1001	GTTTCCGGCTTCGCTTCGGC	-	chr19	54115419	54115438
1002	GGCTTCGCTTCGGCCCCACCC	-	chr19	54115425	54115444
1003	ACGGAGTCTCGTTCTGTTGC	-	chr19	54115924	54115943
1004	TTAGTAGAGACGGGGTTTCC	-	chr19	54116081	54116100
1005	GATCCGCCCCGCCTCGGCCTC	-	chr19	54116139	54116158
1006	CGGAGTTTCGCTCTTGTTGC	-	chr19	54116226	54116245
1007	GATCCACCCGCCTGGGCCTC	-	chr19	54116444	54116463
1008	GCCTTTTACAGCCTGTTTAC	-	chr19	54116503	54116522
1009	CATGGCTCACTCCTGTAATC	-	chr19	54116823	54116842
1010	ATCCCAGCACTTTGGGTGGC	-	chr19	54116840	54116859
1011	GGCGAAAGGATCGCTTGAGC	-	chr19	54116863	54116882
1012	GGCGGGAGAATCGCTTGAGC	-	chr19	54116998	54117017
1013	GGTGGCTTACGCCTGTAATC	-	chr19	54117575	54117594
1014	GGTGGCAGGTGCCTGTAGTC	-	chr19	54117714	54117733
1015	ATGGAGGAGAGGAGAAGGGT	-	chr19	54114067	54114086
1016	TCTGCGTTGTCAAGGAAGCT	-	chr19	54114227	54114246
1017	CTCCACGGGAGATGGCCCTT	-	chr19	54114245	54114264
1018	CAGACCACCTGACACAAACT	-	chr19	54114568	54114587
1019	CACAAACTCCACTTCCGTCT	-	chr19	54114580	54114599
1020	TTCCGTCTCCAGCTCGCTCT	-	chr19	54114592	54114611
1021	CCTCCCTGAGATCCTAGAAT	-	chr19	54114789	54114808
1022	CCTCCCTCAGACCGTAGAAT	-	chr19	54114826	54114845
1023	CCTCCCTCAGACCCAGAAGT	-	chr19	54114861	54114880
1024	TCCTCCCAGGATGCAAGAGT	-	chr19	54114935	54114954
1025	GGATGCAAGAGTCCAGACCT	-	chr19	54114943	54114962
1026	GACCTCCAGACTTTTTCTCT	-	chr19	54114958	54114977
1027	TCTCCAAGGACCCAGGGAGT	-	chr19	54114975	54114994
1028	AGACGCAAGAGTCCTGGCTT	-	chr19	54115015	54115034
1029	AACCTCCTAGTCTGTCAGAT	-	chr19	54115037	54115056
1030	AGTCTGTCAGATCCAGCAGT	-	chr19	54115045	54115064
1031	TCCCTCAGGATGACCCCAGT	-	chr19	54115085	54115104
1032	CACATACCCTTCTCTCTGCT	-	chr19	54115240	54115259
1033	CCCTTCTCTCTGCTCCAATT	-	chr19	54115246	54115265

1034	GCTTCGGCCCCACCCACGT	-	chr19	54115431	54115450
1035	CTCTGCTTAGCAGGATCGGT	-	chr19	54115567	54115586
1036	CGATTCTTCTGCCTCAGCCT	-	chr19	54116005	54116024
1037	GCTGTGATTACAGGCATGAT	-	chr19	54116166	54116185
1038	TATGTTGGCCAGGCTGGTCT	-	chr19	54116405	54116424
1039	AACTCCTGACCTCAGATGAT	-	chr19	54116427	54116446
1040	GGAGTTTTGCAGCTGGGTTT	-	chr19	54117512	54117531

Table 6. Exemplary human genomic region to be targeted.

SEQ ID NO	Human Genomic Region Sequence	Strand	Relative location from TSS
1041	AGATTGCGGGGGGCGGTAGTTTAGGCCAGGGGACCCAAAAC CGGGGGATCCGCACTCACCTACCTGCTCGATCTCCCGGCAGCG CCGACCTAGTGCCTGGTACTTTCTGCGATTTAATTCCCGCTGG CGCCGCCGCCGACCCCGGGCTGCCTCTTCCTCTTCATCTCGCT CCCGGAGCCCTGAGCCGCCAGACCACCTGACACAACTCCA CTTCCGTCTCCAGCTCGCTCTCCAGGATGTGGCCACCAAATAG GGGAGGCAACGCCAACTCTGAGCCTGGCGGCGCTGAGAACTC CTCAAAGCCACGGCTGCCATGGTCCTGAGAGGCAGGGAAAG GCTCAGGGGGCCCTGGATCCTGGACCCCGAGCCCTTCTCCAC TGAACCAGGAGCCAGACCCCAACCCCTCCTCCCTGAGATCCT AGAATCCAGGCCCCAGCCCTCCTCCCTCAGACCGTAGAATC CAGCTCCCAGCCCTCCTCCCTCAGACCCAGAAGTCCAAGTCCG CAACCCACCCTTCGCAGCACCCACAGGGTTCAAGCCCTGACC CCCTCCTCCCAGGATGCAAGAGTCCAGACCTCCAGACTTTTTT TCTCCAAGGACCCAGGGAGTCCAAGCCCCAACCCCTCAACCAG ACGCAAGAGTCCTGGCTTCCAACCTCCTAGTCTGTCAGATCCA GCAGTCCAAACCCCTAACCTTCTCCTCCCTCAGGATGACCCCA GTCCATAAAAGGGTTCTAAGGTAAAGCAGTTGCATGAACTAC AACCCCATCAGACCTCAGCGTAAAGCTCATATGGTTGCAC ACAATGCAGCTGCACTGTTTTCTGGGATTCGCACTTTTTCACA AGGGCTCAGCCACATACCCTTCTCTCTGCTCCAATTCCATCTC CGCGACCTCCGGAAGCCCCGGGCTCAGAGCTTCCGACCTCTT CAATCTGTAGGTTAAGCCGTTGCAAACTACTTGTCCCATCA GGCTCAGCAGCCGAGGACG	+	Upstream

1042	CTTCTCCTCCCTCAGGATGACCCAGTCCATAAAAGGGTTCTA AGGTAAAGCAGTTGCATGAACTACAACCCCCATCAGACCTCA GCGTAAAAGCTCATATGGTTGCACACAATGCAGCTGCACTGTT TTCTGGGATTTCGCACTTTTTTACAAGGGCTCAGCCACATACCC TTCTCTCTGCTCCAATTCCATCTCCGCGACCTCCGGAAGCCCC GGGCCTCAGAGCTTCCGACCTCTTCAATCTGTAGGTTAAGCCG TTCGCAAACTACTTGTCCCATCAGGCTCAGCAGCCGAGGAC G	+	Upstream
1043	CGTCCTCGGCTGCTGAGCCTGATGGGACAAGTAGTTTTGCGAA CGGCTTAACCTACAGATTGAAGAGGTCGGAAGCTCTGAGGCC CGGGGCTTCCGGAGGTCGCGGAGATGGAATTGGAGCAGAGAG AAGGGTATGTGGCTGAGCCCTTGTGAAAAAGTGCGAATCCCA GAAAACAGTGCAGCTGCATTGTGTGCAACCATATGAGCTTTTA CGCTGAGGTCTGATGGGGGTTGTAGTTCATGCAACTGCTTTAC CTTAGAACCCTTTTATGGACTGGGGTCATCCTGAGGGAGGAG AAGGTTAGGGGTTTGGACTGCTGGATCTGACAGACTAGGAGG TTGGAAGCCAGGACTCTTGCGTCTGGTTGAGGGTTGGGGCTTG GACTCCCTGGGTCCCTTGAGAGAAAAAGTCTGGAGGTCTGGA CTCTTGATCCTGGGAGGAGGGGGTCAGGGCTTGAACCCTGT GGGTGCTGCGAAGGGTGGGTTGCGGACTTGGAATTCTGGGTCT TGAGGGAGGAGGGCTGGGAGCTGGATTCTACGGTCTGAGGGA GGAGGGGCTGGGGGCCTGGATTCTAGGATCTCAGGGAGGAGG GGTTGGGGTCTGGGCTCCTGGTTCAGTGGGAGAAGGGGCTGG GGGTCCAGGATCCAGGGCCCCCTGAGCCTTTCCCTGCCTCTCAG GACCATGGCAGCCGTGGGCTTTGAGGAGTTCTCAGCGCCGCC AGGCTCAGAGTTGGCGTTGCCTCCCCTATTTGGTGGCCACATC CTGGAGAGCGAGCTGGAGACGGAAGTGGAGTTTGTGTCAGGT GGTCTGGGCGGCTCAGGGCTCCGGGAGCGAGATGAAGAGGA AGAGGCAGCCCGGGTTCGGCGGCGGCCAGCGGGAATTAA ATCGCAGAAAGTACCAGGCACTAGGTTCGGCGCTGCCGGGAGA TCGAGCAGGTAGGTGAGTGC GGATCCCCCGTTTTTGGGGTCC CCTGGCCTAAACTACCGCCCCCGCAATCT	-	Upstream
1044	CGTCCTCGGCTGCTGAGCCTGATGGGACAAGTAGTTTTGCGAA CGGCTTAACCTACAGATTGAAGAGGTCGGAAGCTCTGAGGCC CGGGGCTTCCGGAGGTCGCGGAGATGGAATTGGAGCAGAGAG AAGGGTATGTGGCTGAGCCCTTGTGAAAAAGTGCGAATCCCA GAAAACAGTGCAGCTGCATTGTGTGCAACCATATGAGCTTTTA	-	Upstream

	CGCTGAGGTCTGATGGGGGTTGTAGTTCATGCAACTGCTTTAC CTTAGAACCCTTTTATGGACTGGGGTCATCCTGAGGGAGGAG AAG		
--	--	--	--

EMBODIMENTS

[0220] The following non-limiting embodiments provide illustrative examples of the invention, but do not limit the scope of the invention.

[0221] Embodiment 1. A system comprising:

a heterologous polypeptide comprising an actuator moiety, wherein the actuator moiety is for binding an endogenous target gene encoding PRPF31 in a cell, to increase expression level of the PRPF31 in the cell, wherein:

(i) the actuator moiety substantially lacks DNA cleavage activity; and/or

(ii) the actuator moiety is coupled to a transcriptional activator,

optionally wherein:

(1) the actuator moiety substantially lacks DNA cleavage activity; and/or

(2) the actuator moiety is coupled to the transcriptional activator; and/or

(3) (i) the actuator moiety substantially lacks DNA cleavage activity and (ii) the actuator moiety is coupled to the transcriptional activator; and/or

(4) the actuator moiety is fused to the transcriptional activator; and/or

(5) the cell is a retinal cell; and/or

(6) the cell is a retinal pigment epithelium (RPE) cell; and/or

(7) the endogenous target gene comprises a non-disease causing allele of the PRPF31; and/or

(8) the actuator moiety is not capable of binding an additional endogenous target gene encoding a mutant allele of the PRPF31; and/or

(9) the target endogenous gene further comprises a disease causing allele of the PRPF31; and/or

(10) the actuator moiety is a deactivated Cas (dCas) protein, further optionally wherein:

(i) a size of the dCas is less than or equal to about 800 amino acids; and/or

(ii) a size of the dCas is less than or equal to about 600 amino acids; and/or

(iii) the dCas protein comprises a polynucleotide sequence exhibiting at least about 90% sequence identity to the polynucleotide sequence selected from Table 1; and/or

(11) the system further comprises a guide nucleic acid capable of forming a complex with the actuator moiety, wherein the complex binds the endogenous target gene,

further optionally wherein the guide nucleic acid comprises a plurality of different guide nucleic acids capable of targeting different polynucleotide sequences of the endogenous target gene; and/or

(12) the heterologous polypeptide and/or the guide nucleic acid is under control of a tissue-specific promoter; and/or

(13) the heterologous polypeptide and/or the guide nucleic acid is under control of a photoreceptor-specific promoter; and/or

(14) the heterologous polypeptide and/or the guide nucleic acid is under control of a constitutive promoter; and/or

(15) the increase in the expression level of the PRPF31 in the cell effects enhanced cilium length and/or incidence, as compared to that in a control cell comprising a mutant allele of the PRPF31 in absence of the heterologous polypeptide and/or the guide nucleic acid; and/or

(16) a portion of the endogenous target gene that is targeted by the actuator moiety and/or the guide nucleic acid is at most about 1,000 nucleobases away from a transcription start site (TSS) of the endogenous target gene; and/or

(17) the guide nucleic acid comprises (i) a scaffold sequence for forming the complex with the actuator moiety and (ii) a spacer sequence exhibiting specific binding to the endogenous target gene, wherein the spacer sequence exhibits at least about 80%, at least about 90%, at least about 95%, or substantially about 100% sequence identity to the polynucleotide sequence selected from Table 5.

[0222] Embodiment 2. One or more polynucleotides encoding the system of Embodiment 1, optionally wherein the one or more polynucleotides comprise a single polynucleotide encoding at least the heterologous polypeptide and the guide nucleic acid,

further optionally wherein:

(i) the single polynucleotide has a size of less than or equal to about 5 kilobases; and/or

(ii) the single polynucleotide has a size of less than or equal to about 4.7 kilobases; and/or

(iii) the single polynucleotide has a size of less than or equal to about 4.2 kilobases; and/or

(iv) the guide nucleic acid comprises the plurality of different guide nucleic acids.

[0223] Embodiment 3. A method comprising administering the system or the one or more polynucleotides of Embodiment 1 or Embodiment 2 to a subject in need thereof,

optionally wherein:

- (1) the administering comprises intravitreal injection or subretinal injection; and/or
- (2) the subject has or is suspected of having retinitis pigmentosa (RP); and/or
- (3) the method further comprises, prior to the administering, determining that the subject has the RP; and/or
- (4) the RP is RP11.

[0224] Embodiment 4. A method comprising:

increasing expression level of an endogenous target gene encoding PRPF31 in a cell, via binding of a heterologous polypeptide comprising an actuator moiety to bind the endogenous target gene, wherein:

- (a) the actuator moiety substantially lacks DNA cleavage activity; and/or
 - (b) the actuator moiety is coupled to a transcriptional activator,
- optionally wherein:
- (1) the actuator moiety substantially lacks DNA cleavage activity; and/or
 - (2) the actuator moiety is coupled to the transcriptional activator; and/or
 - (3) (i) the actuator moiety substantially lacks DNA cleavage activity and (ii) the actuator moiety is coupled to the transcriptional activator; and/or
 - (4) the actuator moiety is fused to the transcriptional activator; and/or
 - (5) the cell is a retinal cell; and/or
 - (6) the cell is a retinal pigment epithelium (RPE) cell; and/or
 - (7) the endogenous target gene comprises a non-disease causing allele of the PRPF31; and/or
 - (8) the actuator moiety is not capable of binding an additional endogenous target gene encoding a mutant allele of the PRPF31; and/or
 - (9) the endogenous target gene further comprises a disease causing allele of the PRPF31; and/or
 - (10) the actuator moiety is a deactivated Cas (dCas) protein, further optionally wherein:
 - (i) a size of the dCas is less than or equal to about 800 amino acids; and/or
 - (ii) a size of the dCas is less than or equal to about 600 amino acids; and/or
 - (iii) the dCas protein comprises a polynucleotide sequence exhibiting at least about 90% sequence identity to the polynucleotide sequence selected from Table 1.
 - (11) the increasing is via action of a complex comprising the actuator moiety and a guide nucleic acid, wherein the complex binds the endogenous target gene,
 - further optionally wherein the guide nucleic acid comprises a plurality of different guide nucleic acids capable of targeting different regions of the endogenous target gene; and/or

(12) the heterologous polypeptide is under control of a tissue-specific promoter; and/or
(13) the heterologous polypeptide is under control of a photoreceptor-specific promoter;
and/or

(14) the heterologous polypeptide and/or the guide nucleic acid is under control of a constitutive promoter; and/or

(15) expression level of the non-disease causing allele of the PRPF31 is enhanced relative to that of a mutant allele of the PRPF31 by at least about 0.1-fold, at least about 1-fold, at least about 5-fold, at least about 10-fold, at least about 50-fold, at least about 100-fold, or at least about 500-fold; and/or

(16) upon the increasing, a phagocytosis level of the cell is reduced by at least about 0.1-fold, at least about 1-fold, at least about 5-fold, at least about 10-fold, at least about 50-fold, at least about 100-fold, or at least about 500-fold, as compared to that of a control cell comprising a mutant allele of the PRPF31 in absence of the heterologous polypeptide and/or the guide nucleic acid; and/or

(17) the increasing effects enhanced cilium length and/or incidence, as compared to that in a control cell comprising a mutant allele of the PRPF31 in absence of the heterologous polypeptide and/or the guide nucleic acid; and/or

(18) upon the increasing, a length of a primary cilia of the cell is longer by at least about 0.1-fold, at least about 1-fold, at least about 5-fold, at least about 10-fold, at least about 50-fold, at least about 100-fold, or at least about 500-fold, as compared to that of a control cell comprising a mutant allele of the PRPF31 in absence of the heterologous polypeptide and/or the guide nucleic acid; and/or

(19) the expression level of the non-disease causing allele of the PRPF31 is substantially the same as that in a cell of a non-diseased cell; and/or

(20) a portion of the endogenous target gene that is targeted by the actuator moiety and/or the guide nucleic acid is at most about 1,000 nucleobases away from a transcription start site (TSS) of the endogenous target gene; and/or

(21) the guide nucleic acid comprises (i) a scaffold sequence for forming the complex with the actuator moiety and (ii) a spacer sequence exhibiting specific binding to the endogenous target gene, wherein the spacer sequence exhibits at least about 80%, at least about 90%, at least about 95%, or substantially about 100% sequence identity to the polynucleotide sequence selected from the group consisting of Table 5.

[0225] It shall be understood that different aspects of the invention can be appreciated individually, collectively, or in combination with each other. Various aspects of the invention

described herein may be applied to any of the particular applications disclosed herein. The compositions of matter disclosed herein in the composition section of the present disclosure may be utilized in the method section including methods of use and production disclosed herein, or vice versa.

[0226] While preferred embodiments of the present invention have been shown and described herein, it will be obvious to those skilled in the art that such embodiments are provided by way of example only. It is not intended that the invention be limited by the specific examples provided within the specification. While the invention has been described with reference to the aforementioned specification, the descriptions and illustrations of the embodiments herein are not meant to be construed in a limiting sense. Numerous variations, changes, and substitutions will now occur to those skilled in the art without departing from the invention. Furthermore, it shall be understood that all aspects of the invention are not limited to the specific depictions, configurations or relative proportions set forth herein which depend upon a variety of conditions and variables. It should be understood that various alternatives to the embodiments of the invention described herein may be employed in practicing the invention. It is therefore contemplated that the invention shall also cover any such alternatives, modifications, variations or equivalents. It is intended that the following claims define the scope of the invention and that methods and structures within the scope of these claims and their equivalents be covered thereby.

CLAIMS

WHAT IS CLAIMED IS:

1. A system comprising:

a heterologous polypeptide comprising an actuator moiety, wherein the actuator moiety is for binding an endogenous target gene encoding PRPF31 in a cell, to increase expression level of the PRPF31 in the cell, wherein:

(i) the actuator moiety substantially lacks DNA cleavage activity; and/or

(ii) the actuator moiety is coupled to a transcriptional activator.

2. The system of claim 1, wherein the actuator moiety substantially lacks DNA cleavage activity.

3. The system of claim 1, wherein the actuator moiety is coupled to the transcriptional activator.

4. The system of claim 1, wherein (i) the actuator moiety substantially lacks DNA cleavage activity and (ii) the actuator moiety is coupled to the transcriptional activator.

5. The system of any one of the preceding claims, wherein the actuator moiety is fused to the transcriptional activator.

6. The system of any one of the preceding claims, wherein the cell is a retinal cell.

7. The system of any one of the preceding claims, wherein the cell is a retinal pigment epithelium (RPE) cell.

8. The system of any one of the preceding claims, wherein the endogenous target gene comprises a non-disease causing allele of the PRPF31.

9. The system of any one of the preceding claims, wherein the actuator moiety is not capable of binding an additional endogenous target gene encoding a mutant allele of the PRPF31.

10. The system of any one of the preceding claims, wherein the target endogenous gene further comprises a disease causing allele of the PRPF31.

11. The system of any one of the preceding claims, wherein the actuator moiety is a deactivated Cas (dCas) protein.

12. The system of claim 11, wherein a size of the dCas is less than or equal to about 800 amino acids.

13. The system of claim 11, wherein a size of the dCas is less than or equal to about 600 amino acids.

14. The system of claim 11, wherein the dCas protein comprises a polynucleotide sequence exhibiting at least about 90% sequence identity to the polynucleotide sequence selected from Table 1.

15. The system of any one of the preceding claims, further comprising a guide nucleic acid capable of forming a complex with the actuator moiety, wherein the complex binds the endogenous target gene.
16. The system of claim 15, wherein the guide nucleic acid comprises a plurality of different guide nucleic acids capable of targeting different polynucleotide sequences of the endogenous target gene.
17. The system of any one of the preceding claims, wherein the heterologous polypeptide and/or the guide nucleic acid is under control of a tissue-specific promoter.
18. The system of any one of the preceding claims, wherein the heterologous polypeptide and/or the guide nucleic acid is under control of a photoreceptor-specific promoter.
19. The system of any one of the preceding claims, wherein the heterologous polypeptide and/or the guide nucleic acid is under control of a constitutive promoter.
20. The system of any one of the preceding claims, wherein the increase in the expression level of the PRPF31 in the cell effects enhanced cilium length and/or incidence, as compared to that in a control cell comprising a mutant allele of the PRPF31 in absence of the heterologous polypeptide and/or the guide nucleic acid.
21. The system of any one of the preceding claims, wherein a portion of the endogenous target gene that is targeted by the actuator moiety and/or the guide nucleic acid is at most about 1,000 nucleobases away from a transcription start site (TSS) of the endogenous target gene.
22. The system of any one of the preceding claims, wherein the guide nucleic acid comprises (i) a scaffold sequence for forming the complex with the actuator moiety and (ii) a spacer sequence exhibiting specific binding to the endogenous target gene, wherein the spacer sequence exhibits at least about 80%, at least about 90%, at least about 95%, or substantially about 100% sequence identity to the polynucleotide sequence selected from Table 5.
23. One or more polynucleotides encoding the system of any one of the preceding claims.
24. The one or more polynucleotides of claim 23, wherein the one or more polynucleotides comprise a single polynucleotide encoding at least the heterologous polypeptide and the guide nucleic acid.
25. The one or more polynucleotides of claim 24, wherein the single polynucleotide has a size of less than or equal to about 5 kilobases.
26. The one or more polynucleotides of claim 24, wherein the single polynucleotide has a size of less than or equal to about 4.7 kilobases.
27. The one or more polynucleotides of claim 24, wherein the single polynucleotide has a size of less than or equal to about 4.2 kilobases.

28. The one or more polynucleotides of claim 24, wherein the guide nucleic acid comprises the plurality of different guide nucleic acids.
29. A method comprising administering the system or the one or more polynucleotides of any one of the preceding claims to a subject in need thereof.
30. The method of claim 29, wherein the administering comprises intravitreal injection or subretinal injection.
31. The method of any one of the preceding claims, wherein the subject has or is suspected of having retinitis pigmentosa (RP).
32. The method of any one of the preceding claims, further comprising, prior to the administering, determining that the subject has the RP.
33. The method of any one of the preceding claims, wherein the RP is RP11.
34. A method comprising:
- increasing expression level of an endogenous target gene encoding PRPF31 in a cell, via binding of a heterologous polypeptide comprising an actuator moiety to bind the endogenous target gene, wherein:
 - (a) the actuator moiety substantially lacks DNA cleavage activity; and/or
 - (b) the actuator moiety is coupled to a transcriptional activator.
35. The method of claim 34, wherein the actuator moiety substantially lacks DNA cleavage activity.
36. The method of claim 34, wherein the actuator moiety is coupled to the transcriptional activator.
37. The method of claim 34, wherein (i) the actuator moiety substantially lacks DNA cleavage activity and (ii) the actuator moiety is coupled to the transcriptional activator.
38. The method of any one of the preceding claims, wherein the actuator moiety is fused to the transcriptional activator.
39. The method of any one of the preceding claims, wherein the cell is a retinal cell.
40. The method of any one of the preceding claims, wherein the cell is a retinal pigment epithelium (RPE) cell.
41. The method of any one of the preceding claims, wherein the endogenous target gene comprises a non-disease causing allele of the PRPF31.
42. The method of any one of the preceding claims, wherein the actuator moiety is not capable of binding an additional endogenous target gene encoding a mutant allele of the PRPF31.
43. The method of any one of the preceding claims, wherein the endogenous target gene further comprises a disease causing allele of the PRPF31.

44. The method of any one of the preceding claims, wherein the actuator moiety is a deactivated Cas (dCas) protein.
45. The method of claim 44, wherein a size of the dCas is less than or equal to about 800 amino acids.
46. The method of claim 44, wherein a size of the dCas is less than or equal to about 600 amino acids.
47. The method of claim 44, wherein the dCas protein comprises a polynucleotide sequence exhibiting at least about 90% sequence identity to the polynucleotide sequence selected from Table 1.
48. The method of any one of the preceding claims, wherein the increasing is via action of a complex comprising the actuator moiety and a guide nucleic acid, wherein the complex binds the endogenous target gene.
49. The method of claim 48, wherein the guide nucleic acid comprises a plurality of different guide nucleic acids capable of targeting different regions of the endogenous target gene.
50. The method of any one of the preceding claims, wherein the heterologous polypeptide is under control of a tissue-specific promoter.
51. The method of any one of the preceding claims, wherein the heterologous polypeptide is under control of a photoreceptor-specific promoter.
52. The method of any one of the preceding claims, wherein the heterologous polypeptide and/or the guide nucleic acid is under control of a constitutive promoter.
53. The method of any one of the preceding claims, wherein expression level of the non-disease causing allele of the PRPF31 is enhanced relative to that of a mutant allele of the PRPF31 by at least about 0.1-fold, at least about 1-fold, at least about 5-fold, at least about 10-fold, at least about 50-fold, at least about 100-fold, or at least about 500-fold.
54. The method of any one of the preceding claims, wherein, upon the increasing, a phagocytosis level of the cell is reduced by at least about 0.1-fold, at least about 1-fold, at least about 5-fold, at least about 10-fold, at least about 50-fold, at least about 100-fold, or at least about 500-fold, as compared to that of a control cell comprising a mutant allele of the PRPF31 in absence of the heterologous polypeptide and/or the guide nucleic acid.
55. The method of any one of the preceding claims, wherein the increasing effects enhanced cilium length and/or incidence, as compared to that in a control cell comprising a mutant allele of the PRPF31 in absence of the heterologous polypeptide and/or the guide nucleic acid.

56. The method of any one of the preceding claims, wherein, upon the increasing, a length of a primary cilia of the cell is longer by at least about 0.1-fold, at least about 1-fold, at least about 5-fold, at least about 10-fold, at least about 50-fold, at least about 100-fold, or at least about 500-fold, as compared to that of a control cell comprising a mutant allele of the PRPF31 in absence of the heterologous polypeptide and/or the guide nucleic acid.
57. The method of any one of the preceding claims, wherein the expression level of the non-disease causing allele of the PRPF31 is substantially the same as that in a cell of a non-diseased cell.
58. The method of any one of the preceding claims, wherein a portion of the endogenous target gene that is targeted by the actuator moiety and/or the guide nucleic acid is at most about 1,000 nucleobases away from a transcription start site (TSS) of the endogenous target gene.
59. The method of any one of the preceding claims, wherein the guide nucleic acid comprises (i) a scaffold sequence for forming the complex with the actuator moiety and (ii) a spacer sequence exhibiting specific binding to the endogenous target gene, wherein the spacer sequence exhibits at least about 80%, at least about 90%, at least about 95%, or substantially about 100% sequence identity to the polynucleotide sequence selected from the group consisting of Table 5.

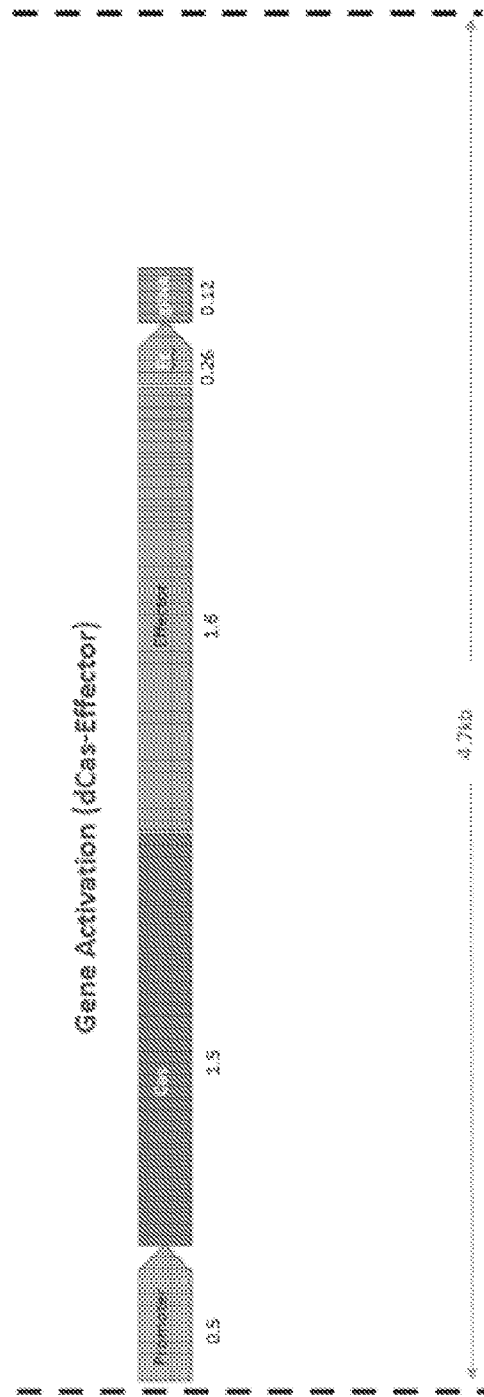


FIG. 1

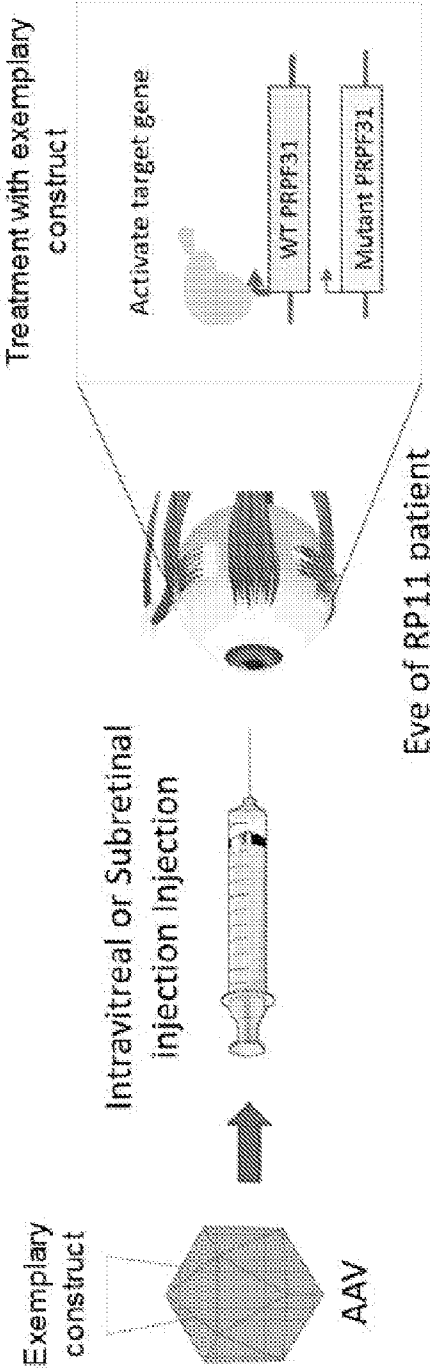


FIG. 2

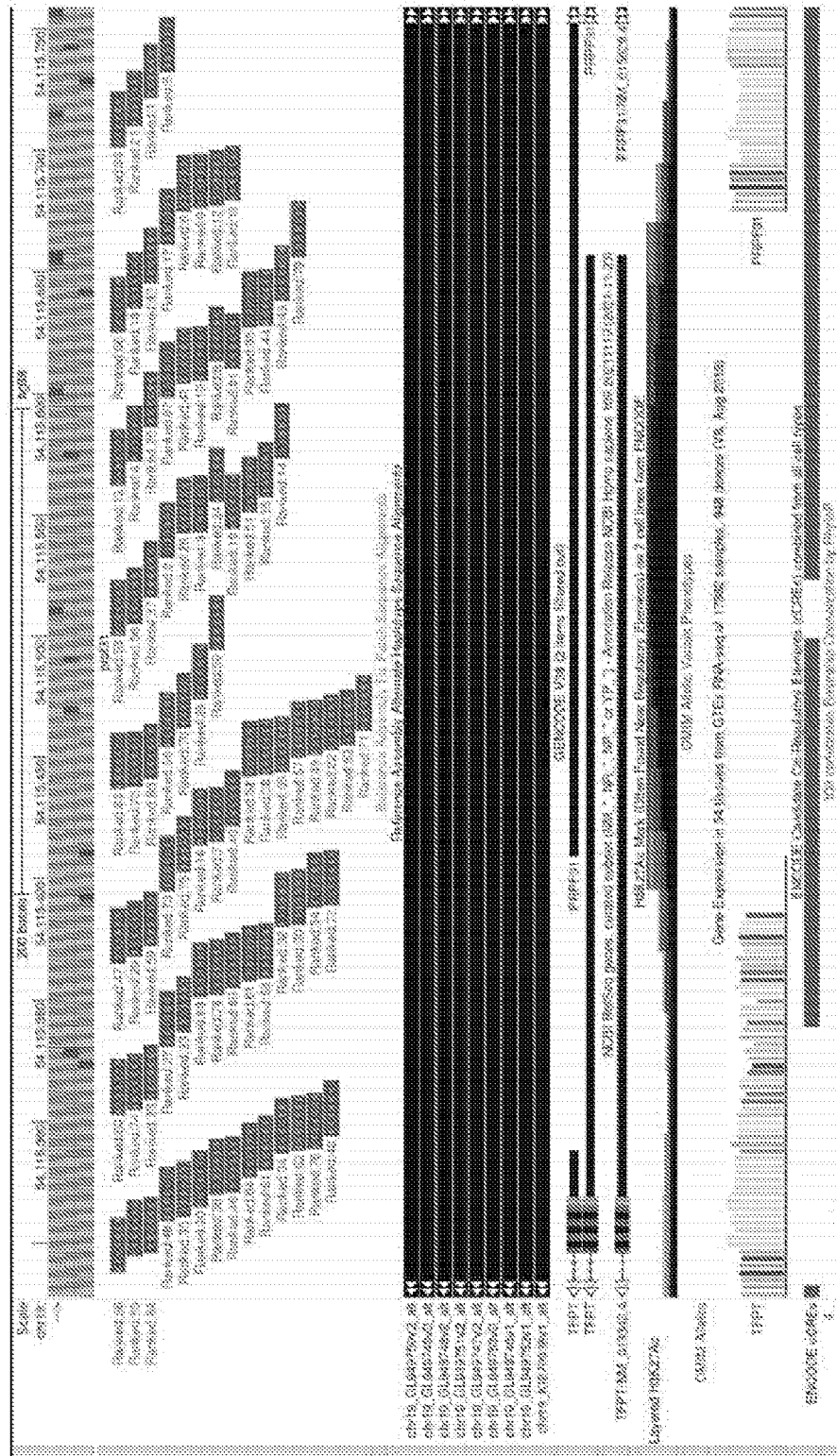
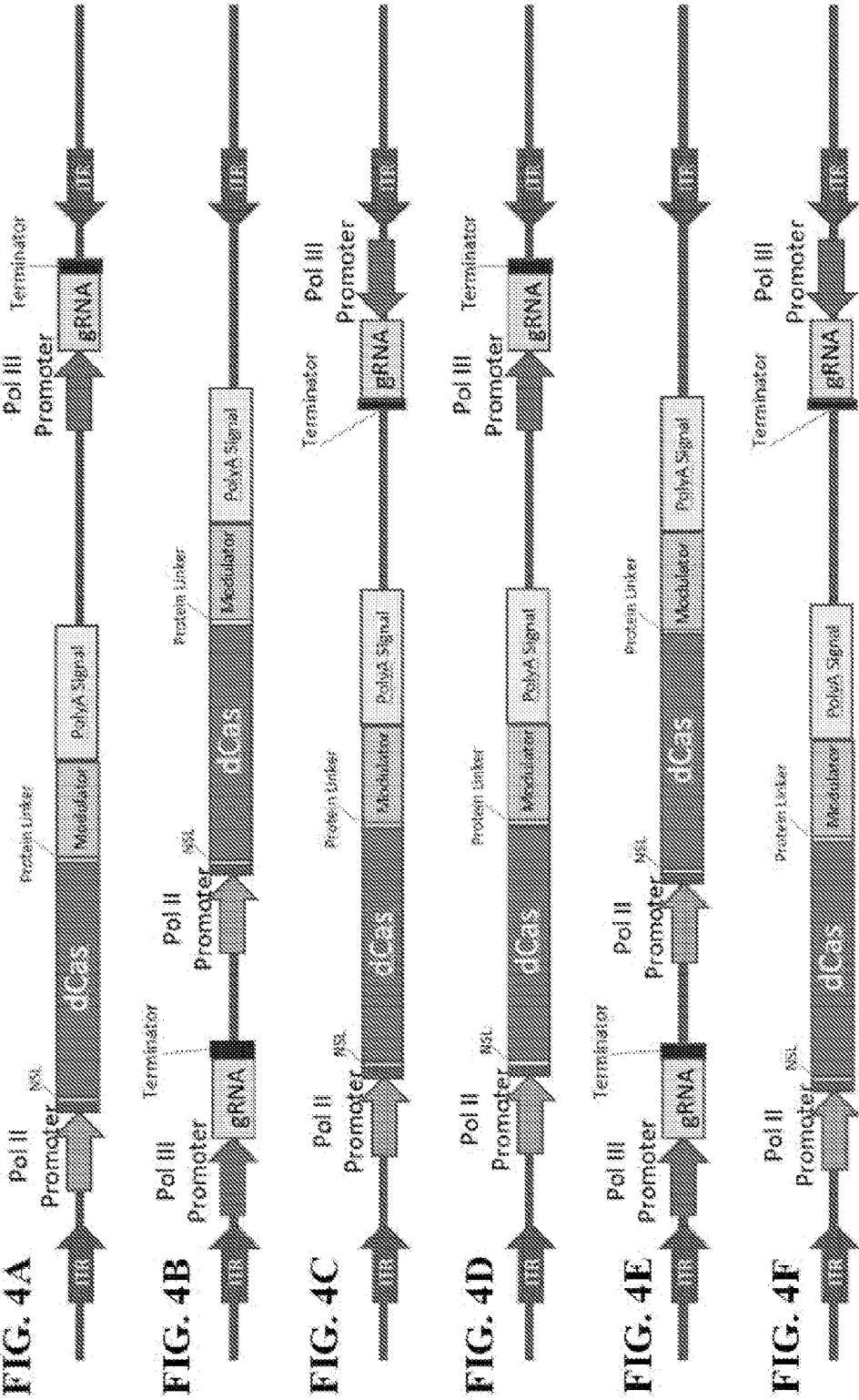


FIG. 3



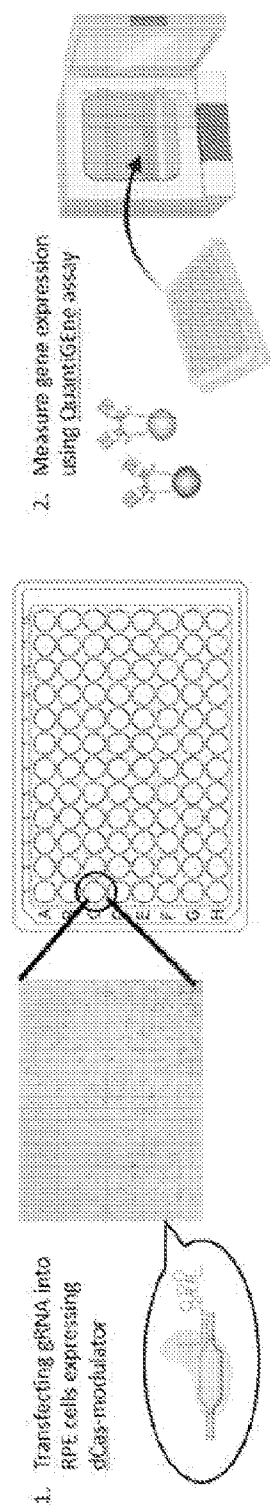


FIG. 5A

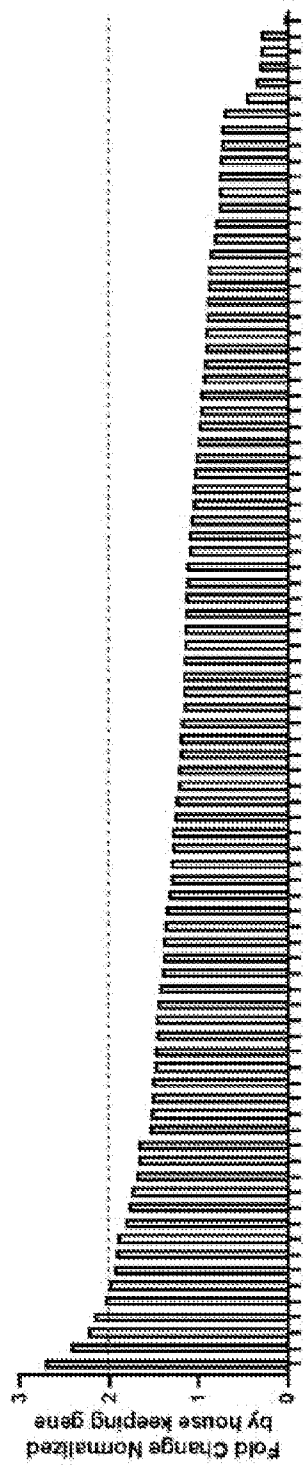


FIG. 5B

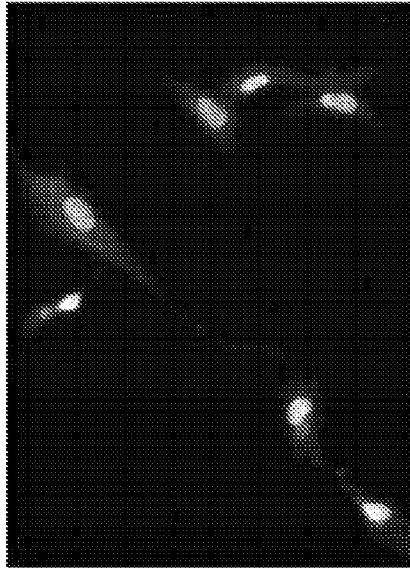


FIG. 6B

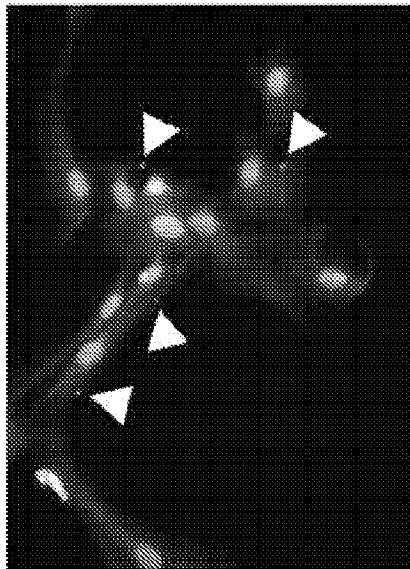


FIG. 6A

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2023/064004

A. CLASSIFICATION OF SUBJECT MATTER

**C12N 15/113 (2010.01) A61K 48/00 (2006.01) A61P 27/02 (2006.01) C07K 14/47 (2006.01) C12N 5/079 (2010.01)
C12N 9/22 (2006.01) C12N 15/62 (2006.01) C12N 15/63 (2006.01)**

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)

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

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

STN (MEDLINE, CAPLUS, BIOSIS, EMBASE databases): PRPF31, SNRNP61, retinitis pigmentosa, ocular disease, Cas, dCas, ZFN, TALEN, transcriptional activators, fused, spacer, CRISPRa, base editing and like terms; **EPOQUE** (PATENW English language databases): C12N9/22, C12N2310/20, C12N15/62, PRPF31, SNRNP61, retinitis pigmentosa, ocular disease, Cas, dCas, ZFN, TALEN, transcriptional activators, and like terms; **GenomeQuest**: SEQ ID NOs: 1-201 and 786-1040; **Google**: dCas, fusion, VPR, PRPF31; **Automated Preliminary Search Tool** (IP Australia's internal databases): Applicant and inventors names.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
	Documents are listed in the continuation of Box C	

☒ 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	"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
"D" document cited by the applicant in the international application	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search
26 May 2023

Date of mailing of the international search report
26 May 2023

Name and mailing address of the ISA/AU

AUSTRALIAN PATENT OFFICE
PO BOX 200, WODEN ACT 2606, AUSTRALIA
Email address: pct@ipaustalia.gov.au

Authorised officer

Joan Holgate
AUSTRALIAN PATENT OFFICE
(ISO 9001 Quality Certified Service)
Telephone No. +61 2 6283 2671

INTERNATIONAL SEARCH REPORT		International application No.
C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		PCT/US2023/064004
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
L	Buskin A, et al. Disrupted alternative splicing for genes implicated in splicing and ciliogenesis causes PRPF31 retinitis pigmentosa. Nature communications. 2018 Oct 12;9(1):4234. Lack of unity	
X	US 20190351074 A1 (The Regents of the University of California) 21 November 2019 [0007]; Claims 1, 40; Table 1; Figures 1a-1c	1-59
X	WO 2019210305 A1 (The trustees of the Columbia University in the city of New York) 31 October 2019 page 3, lines 5-30; page 23, lines 14-16, 23-30; page 33, lines 30-31; page 25, lines 5-14; page 26, lines 10-22; page 26, lines 18-19; page 57, lines 1-23; Claims 6, 19, 32	1-59
A	US 10934536 B2 (Pioneer Hi-Bred International, Inc.) 02 March 2021 SEQ ID NO: 38	14, 47
A	WO 2022020393 A1 (Mammoth Biosciences, Inc.) 27 January 2022 SEQ ID NOS: 3, 10	14, 47

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2023/064004

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.
 - b. ☒ furnished subsequent to the international filing date for the purposes of international search (Rule 13*ter*.1(a)).
☒ accompanied by a statement to the effect that the sequence listing does not go beyond the disclosure in the international application as filed.
2. ☐ With regard to any nucleotide and/or amino acid sequence disclosed in the international application, this report has been established to the extent that a meaningful search could be carried out without a WIPO Standard ST.26 compliant sequence listing.
3. Additional comments:

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US2023/064004

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:
the subject matter listed in Rule 39 on which, under Article 17(2)(a)(i), an international search is not required to be carried out, including
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.:
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:

See Supplemental Box for Details

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.
2. ☒ As all searchable claims could be searched without effort justifying additional fees, this Authority did not invite payment of additional fees.
3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

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

Supplemental Box**Continuation of: Box III**

This International Application does not comply with the requirements of unity of invention because it does not relate to one invention or to a group of inventions so linked as to form a single general inventive concept.

This Authority has found that there are different inventions based on the following features that separate the claims into distinct groups:

Invention 1: Claims 1-2, 4-35 and 37-39 are directed to the feature of an actuator moiety which substantially lacks DNA cleavage activity;

Invention 2: Claims 1, 3-10, 15-34, 36-43 and 48-59 are directed to the feature of an actuator moiety which is coupled to a transcriptional activator.

PCT Rule 13.2, first sentence, states that unity of invention is only fulfilled when there is a technical relationship among the claimed inventions involving one or more of the same or corresponding special technical features. PCT Rule 13.2, second sentence, defines a special technical feature as a feature which makes a contribution over the prior art.

When there is no special technical feature common to all the claimed inventions there is no unity of invention.

In the above groups of claims, the identified features may have the potential to make a contribution over the prior art but are not common to all the claimed inventions and therefore cannot provide the required technical relationship. The only feature common to all of the claimed inventions and which provides a technical relationship among them is an actuator moiety for binding an endogenous target gene encoding PRPF31 in a cell and increasing the expression level of the PRPF31. However this feature does not make a contribution over the prior art because it is disclosed in:

D1 discloses a CRISPR-Cas9 system for the correction of a PRPF31 mutation which comprises Cas9 (an actuator moiety) which binds the endogenous PRPF31 gene in RP11VS-induced pluripotent stem cells (iPSCs) and increases the expression level of PRPF31 (CRISPR-Cas9 correction of PRPF31 mutation in the RP11VS, page 15; Figures 7a-7m).

Therefore in the light of this document this common feature cannot be a special technical feature. Therefore there is no special technical feature common to all the claimed inventions and the requirements for unity of invention are consequently not satisfied *a posteriori*.

All claims have been searched and examined.

INTERNATIONAL SEARCH REPORT		International application No.	
Information on patent family members		PCT/US2023/064004	
This Annex lists known patent family members relating to the patent documents cited in the above-mentioned international search report. The Australian Patent Office is in no way liable for these particulars which are merely given for the purpose of information.			
Patent Document/s Cited in Search Report		Patent Family Member/s	
Publication Number	Publication Date	Publication Number	Publication Date
US 20190351074 A1	21 November 2019	US 2019351074 A1	21 Nov 2019
		AU 2018218280 A1	29 Aug 2019
		AU 2018218280 A2	31 Oct 2019
		CN 110612113 A	24 Dec 2019
		EP 3579858 A1	18 Dec 2019
		JP 2020506948 A	05 Mar 2020
		JP 2023058644 A	25 Apr 2023
		WO 2018148256 A1	16 Aug 2018
WO 2019210305 A1	31 October 2019	WO 2019210305 A1	31 Oct 2019
		EP 3810766 A1	28 Apr 2021
		US 2021054372 A1	25 Feb 2021
US 10934536 B2	02 March 2021	US 2020190494 A1	18 Jun 2020
		US 10934536 B2	02 Mar 2021
		AU 2019398351 A1	03 Jun 2021
		BR 112021011372 A2	31 Aug 2021
		CA 3117228 A1	18 Jun 2020
		CN 113166744 A	23 Jul 2021
		EP 3894550 A2	20 Oct 2021
		IL 283853 A	29 Jul 2021
		JP 2022514493 A	14 Feb 2022
		KR 20210104068 A	24 Aug 2021
		US 2021139874 A1	13 May 2021
		US 2021163908 A1	03 Jun 2021
		US 2022073890 A1	10 Mar 2022
		US 2023119655 A1	20 Apr 2023
		WO 2020123887 A2	18 Jun 2020
WO 2022020393 A1	27 January 2022	WO 2022020393 A1	27 Jan 2022
End of Annex			
Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.			

Due to data integration issues this family listing may not include 10 digit Australian applications filed since May 2001.

Form PCT/ISA/210 (Family Annex)(July 2019)